

STIC-Biotech/ChemLib

160642 160624

From: Ibrahim, Medina A.
Sent: Tuesday, July 26, 2005 6:18 PM
To: STIC-Biotech/ChemLib
Subject: 09/765, 555

Please search the following:

Please search the following:

1. DNA encoding SEQ ID NO: 14- 18. Thanks

Medina A. Ibrahim
Patent Examiner, GAU-1638
transgenic plants and plant breeding
Remsen-2B07
mailbox-2C18
(571) 272-0797

STAFF USE ONLY

Searcher: _____
Searcher Phone: 2- _____
Date Searcher Picked up: 8/8/05
Date Completed: 8/8/05
Searcher Prep/Rev. Time: _____
Online Time: _____

Type of Search

NA#: 5 AA#: _____
Interference: _____ SPDI: _____
S/L: _____ Oligomer: _____
Encode/Transl: _____
Structure#: _____ Text: _____
Inventor: _____ Litigation: _____

Vendors and cost where applicable

STN: _____
DIALOG: _____
QUESTEL/ORBIT: _____
LEXIS/NEXIS: _____
SEQUENCE SYSTEM: pat
WWW/Internet: _____
Other(Specify): _____

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 6743.2 Seconds
(without alignments)
18627.964 Million cell updates/sec

Title: US-09-765-555B-16
Perfect score: 3300
Sequence: 1 ccgacacatgaatgtgc.....acgacgttcgcgactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	778	23.6	794	5	BQ751936	BQ751936 EST632499
2	770.6	23.4	779	7	CK394523	CK394523 hggada4D0
3	768	23.3	769	7	CK118014	CK118014 218n06.p1
4	762	23.1	793	5	BQ751655	BQ751655 EST632218
5	755.4	22.9	806	5	BQ751220	BQ751220 EST631783
6	684.8	20.8	689	6	CB863541	CB863541 HH04A08Y
7	676.4	20.5	712	1	AL038548	AL038548 DKF2P566E
8	676	20.5	710	8	AY080106	AY080106 AY080106
9	647.4	19.6	660	1	AL044483	AL044483 DKF2P434I
10	641	19.4	641	4	BJ063872	BJ063872 BJ063872
11	615.8	18.7	624	6	CB863814	CB863814 HH07B20Y
12	612	18.5	794	1	AL045353	AL045353 DKF2P434B
13	603	18.3	752	9	CL655575	CL655575 PRI0124a
14	589.8	17.9	608	8	AY080094	AY080094 AY080094
15	586.4	17.8	589	1	AL037742	AL037742 DKF2P564C
16	581.2	17.6	601	1	AV594538	AV594538 AV594538
17	579	17.5	579	1	AL043868	AL043868 DKF2P434C
18	570	17.3	782	7	CK394246	CK394246 hggada1B1
19	569	17.2	774	1	AL045337	AL045337 DKF2P434A
20	563.2	17.1	585	9	CL658017	CL658017 PRI0130b
21	562.6	17.0	723	3	AL039416	AL039416 DKF2P434I
22	549.4	16.6	753	1	AL045341	AL045341 DKF2P434B
23	546.8	16.6	782	1	AL037051	AL037051 DKF2P564I
24	545.4	16.5	547	6	CA895956	CA895956 B0195F10-

C 25	545	16.5	648	7	CK394349	CK394349 hggada2D0
C 26	540.2	16.4	569	1	AL039077	AL039077 DKF2P566G
C 27	534.4	16.2	644	6	CB862524	CB862524 HH04B08U
C 28	534	16.2	734	1	AL048427	AL048427 DKF2P588H
C 29	529	16.0	740	1	AL042909	AL042909 DKF2P434J
C 30	528	16.0	720	6	CA881974	CA881974 K0994C10-
C 31	523.2	15.9	528	1	AL039811	AL039811 DKF2P566O
C 32	519.4	15.7	521	6	CA890032	CA890032 B0136F03-
C 33	519	15.7	607	6	CB862099	CB862099 HH06A02Y
C 34	516.4	15.6	518	6	CA886082	CA886082 B0124F04-
C 35	510.2	15.5	686	1	AL044407	AL044407 DKF2P434D
C 36	509.4	15.4	511	6	CA895436	CA895436 B0192A07-
C 37	498	15.1	498	1	AL039076	AL039076 DKF2P566G
C 38	494.8	15.0	1025	1	AL038025	AL038025 DKF2P566C
C 39	487.6	14.8	498	7	CK394397	CK394397 hggada2H1
C 40	477.6	14.5	491	6	CB862715	CB862715 HH03F14U
C 41	465.8	14.1	679	1	AL039128	AL039128 DKF2P566K
C 42	462.8	14.0	490	1	AL039589	AL039589 DKF2P434D
C 43	458	13.9	458	6	CA887583	CA887583 B0137H04-
C 44	456.4	13.8	471	1	AL039649	AL039649 DKF2P434G
C 45	449.2	13.6	612	6	CB863409	CB863409 HH04K05W

ALIGNMENTS

RESULT 1
BQ751936 794 bp mRNA linear EST 18-JUL-2002
LOCUS EST632499 DSCT Colletotrichum trifolii cDNA clone pDSC78-67, mRNA
DEFINITION sequence.
ACCESSION BQ751936
VERSION BQ751936.1 GI:21907341
KEYWORDS EST.
SOURCE Colletotrichum trifolii
ORGANISM Colletotrichum trifolii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
mitosporic Phyllachoraceae; Colletotrichum.
REFERENCE 1 (bases 1 to 794)
AUTHORS Samac D.A., Dickman, M., Town, C.D., Van Aken, S., Utterback, T.,
Cheung, F. and Fraser, C.M.
ESTS from mycelia of Colletotrichum trifolii race 1
Other ESTs: EST632498
Contact: Deborah A. Samac
Department of Plant Pathology
University of Minnesota
495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
Tel: 612 625 1243
Fax: 651 649 5058
Email: debbys@puccini.crl.umn.edu
TIGR sequence name: MTSAH67TV More information is available at:
www.medicago.org
Seq primer: lgtA AtA CGA Ctc Act AtA 999 C).
Location/Qualifiers
1..794
/organism="Colletotrichum trifolii"
/mol_type="mRNA"
/strain="race 1"
/db_xref="taxon:5466"
/clone="pDSC78-67"
/tissue_type="mycelia"
/dev_stage="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2%glucose)."
/lab_host="DH5alpha"
/clone_lib="DSCT"
/note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; Isolate: 28p2 ; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gill from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce.

FEATURES
source

E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN

Query Match 23.6%; Score 778; DB 5; Length 794;
Best Local Similarity 98.7%; Pred. No. 1.3e-217;
Matches 784; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 3 GACACATCGAATGGTGCAGAAACCTTTCGCGTATGATGATAGCGCGGAGAGAGT 62
DB 1 GACACTTTCGAATGGCGCAAAACCTTTCGCGTATGATGATAGCGCGGAGAGAGT 60

QY 63 CAATTTCAGGTCGTAATGTAACAGTAAAGTATACGATGTCGAGATATCCCGT 122
DB 61 CAATTTCAGGTCGTAATGTAACAGTAAAGTATACGATGTCGAGATATCCCGT 120

QY 123 GTCTCTTATCAGACCGTTTCCCGCGTGTGAACAGGCGCAGCTTCTGCGAAACG 182
DB 121 GTCTCTTATCAGACCGTTTCCCGCGTGTGAACAGGCGCAGCTTCTGCGAAACG 180

QY 183 CGGGAAGAGTGAAGCGCGGATGGCGAGCTGAATTAATTCACACCGCGTGSCAAA 242
DB 181 CGGGAAGAGTGAAGCGCGGATGGCGAGCTGAATTAATTCACACCGCGTGSCAAA 240

QY 243 CAACTGGCGGCAACAGTCGTGATGCGGTCGCGTGCACCTCAGTCTGGCCCTGAC 302
DB 241 CAACTGGCGGCAACAGTCGTGATGCGGTCGCGTGCACCTCAGTCTGGCCCTGAC 300

QY 303 GCGCGTCGCAAAATGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCGT 362
DB 301 GCGCGTCGCAAAATGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCGT 360

QY 363 GTGGTGTGATGATGAGAGAGCGCGTGCAGCTGTAAAGCGCGGTGCACAACTTT 422
DB 361 GTGGTGTGATGATGAGAGAGCGCGTGCAGCTGTAAAGCGCGGTGCACAACTTT 420

QY 423 CTCGCGCAACCGTCAGTGGGCTGATCAATTAATCTCGCGTGGATGACAGGATGCCATT 482
DB 421 CTCGCGCAACCGTCAGTGGGCTGATCAATTAATCTCGCGTGGATGACAGGATGCCATT 480

QY 483 GCTGTGGAAGCTGCTGCACTAAATGTTCCGCGTATTTCTTGATGTCCTGACAGACA 542
DB 481 GCTGTGGAAGCTGCTGCACTAAATGTTCCGCGTATTTCTTGATGTCCTGACAGACA 540

QY 543 CCCATCAACAGTATTTCTCCATGAAGACGCTAGCGACTGGGCGGTGAGCATCTG 602
DB 541 CCCATCAACAGTATTTCTCCATGAAGACGCTAGCGACTGGGCGGTGAGCATCTG 600

QY 603 GTCGATTTGGTTCACAGCAATCGCGTGTAGCGGCGCCATTAAGTCTGCTCGCGG 662
DB 601 GTCGATTTGGTTCACAGCAATCGCGTGTAGCGGCGCCATTAAGTCTGCTCGCGG 660

QY 663 CGTCTGCGTCTGGCTGGCTGAATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
DB 661 CGTCTGCGTCTGGCTGGCTGGCGATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 720

QY 723 GAAACGGAAGCGACTGGAGTGCCATGTCGGTGTTCACAAACCATGCAATGCTGAAT 782
DB 721 GAAACGGAAGCGACTGGAGTGCCATGTCGGTGTTCACAAACCATGCAATGCTGAAT 780

QY 783 GAGGGCATCGTTCC 796
DB 781 GAGGGCATCGTTCC 794

RESULT 2
CK394523/c
LOCUS hggada4D08 Gland Cell Amplified cDNA Library Heterodera glycines
DEFINITION

CDNA, mRNA sequence.
CK394523
CK394523.1 GI:40389794
EST.
Heterodera glycines
Heterodera glycines
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera.
1 (bases 1 to 779)
Wang,X., Allen,R., Gao,B., Goellner,M., Maier,T., Baum,T.,
Hussey,R. and Davis,E.
Submission: Wang, Allen, Gao, Goellner, Maier, Baum, Hussey, Davis
Unpublished (2003)
Contact: Tom Maier
Department of Plant Pathology, Baum Lab
Iowa State University
351 Bessey Hall, Ames, IA 50011, USA
Tel: 515-294-8854
Fax: 515-294-9420
Email: tmaier@iastate.edu
Heterodera glycines Gland Cell Amplified cDNA Library, single pass
sequence.

FEATURES
source
1..779
/organism="Heterodera glycines"
/mol_type="mRNA"
/db_xref="taxon:51029"
/cell_type="gland"
/dev_stage="mixed parasitic juvenile"
/clone_lib="Gland Cell Amplified cDNA Library"
/note="Organ: gland cell; Vector: pSportII"

ORIGIN

Query Match 23.4%; Score 770.6; DB 7; Length 779;
Best Local Similarity 99.4%; Pred. No. 2e-215;
Matches 773; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 87 CCAGTAACGTTATACGATGTCGAGAGTATGCCGCTGCTCTTATCAGACCGTTCCCGC 146
DB 779 CCAGTAACGTTATACGATGTCGAGAGTATGCCGCTGCTCTTATCAGACCGTTCCCGC 720

QY 147 GTGCTGAACAGCCAGCCAGCCAGCTTCTGCGAAACCGCGGAAAAAGTGAAGCGCGATG 206
DB 719 GTGCTGAACAGCCAGCCAGCCAGCTTCTGCGAAACCGCGGAAAAAGTGAAGCGCGATG 660

QY 207 GCGGAGCTGAATTACATTCCTCAACCGCGTGGCAACAACAACTGGCGGGCAACAGTCGTTG 266
DB 659 GCGGAGCTGAATTACATTCCTCAACCGCGTGGCAACAACAACTGGCGGGCAACAGTCGTTG 600

QY 267 CTGATTGGCGTTGCGCACCTCCAGTCTGGCCCTGCGACCGCGCTCGCAAAATTTGTCGGCGG 326
DB 599 CTGATTGGCGTTGCGCACCTCCAGTCTGGCCCTGCGACCGCGCTCGCAAAATTTGTCGGCGG 540

QY 327 ATTAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTGTGATGTAGAACGAAGC 386
DB 539 ATTAATCTCGCGCCGATCAACTGGGTGCCAGCGTGGTGGTGTGATGTAGAACGAAGC 480

QY 387 GCGCTGGAAGCTGTAAAGCGGCGGTGCACATCTTCTCGCGCAACCGGTGAGTGGGCTG 446
DB 479 GCGCTGGAAGCTGTAAAGCGGCGGTGCACATCTTCTCGCGCAACCGGTGAGTGGGCTG 420

QY 447 ATCATTAATCTCGCGTGGATGACAGGATGCAATGCTGTGGAAGCTGCCCTGCACTAAT 506
DB 419 ATCATTAATCTCGCGTGGATGACAGGATGCAATGCTGTGGAAGCTGCCCTGCACTAAT 360

QY 507 GTTCGCGGTTATTTCTTGATGTCGTGACACAGACACCCATCAACAGTATTTTCTCC 566
DB 359 GTTCGCGGTTATTTCTTGATGTCGTGACACAGACACCCATCAACAGTATTTTCTCC 300

QY 567 CATGAAGACGCTACCGACTGGCGGTGGAGCATCTGGTCGATTCGGTCCACAGCAATC 626
DB 299 CATGAAGACGCTACCGACTGGCGGTGGAGCATCTGGTCGATTCGGTCCACAGCAATC 240

COMMENT

Other ESTs: EST632217
Contact: Deborah A. Samac
Department of Plant Pathology
University of Minnesota
495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
Tel: 612 625 1243
Fax: 651 649 5058
Email: debbys@puccini.crl.umn.edu
TIGR sequence name: MTSAF59rv More information is available at:
www.medicago.org
Seq primer: (gtA AtA CgA Ctc ACt AtA qqq C).

FEATURES

source

```

location/qualifiers
1..793
/organism="Colletotrichum trifolii"
/mol_type="mRNA"
/strain="race 1"
/db_xref="taxon:5466"
/clones="pDSCF6-59"
/tissue_type="mycelia"
/dev_stage="young, actively growing
inoculation) grown in liquid culture
containing 2%glucose)."
/lab_host="D54alpha"
/clone_lib="BSCt"
/notes="vector: pBluescript SK-; Site
ECORI; isolate: 2sp2 ; cDNA was prepared
Stratagene RNA The cDNA was ligated into
Stratagene and packaged using Gigapack
An aliquot of the amplified library
E. coli J0390 and phage DNA was purified
lyate. The cDNA inserts were gel purified
digestion and ligated into pBluescript
the ligation were used to transform
were plated onto medium with X-gal for
recombinants."

```

ORIGIN

Query Match	23.1%	Score 762;	DB 5;	Length 793;
Best Local Similarity	99.1%;	Pred. No. 7e-213;		
Matches 777;	Conservative 0;	Mismatches 5;	Indels 2;	Gaps 1
QY	3	GACACCATCGAATGCTGCAAAACCTTTTCGCGGTATGCATCATAGCGCCCGGAAGAGT	62	
Db	1	GACATCTTCGAATGCGCGCAAAACCTTTTCGCGGTATGCATCATAGCGCCCGGAAGAGT	60	
QY	63	CAATTACGGGTGGTGAATGTGAAACCAAGTAAACGTTATACGATGTGCAGATGATGCCGCT	122	
Db	61	CAATTACGGGTGGTGAATGTGAAACCAAGTAAACGTTATACGATGTGCAGATGATGCCGCT	120	
QY	123	GTCTCTTATCAGACCGTTTCCCGCGTGTGTAACACGAGCCAGCCAGCTTTCTCGGAAACG	182	
Db	121	GTCTCTTATCAGACCGTTTCCCGCGTGTGTAACACGAGCCAGCCAGCTTTCTCGGAAACG	180	
QY	193	CGGGAAAAAGTGGAAACGCGGATCGGCGAGCTGAATTACATTTCCCAACCGCGTGGCACA	242	
Db	181	CGGGAAAAAGTGGAAACGCGGATCGGCGAGCTGAATTACATTTCCCAACCGCGTGGCACA	240	
QY	243	CAACTGGCGGGCAACACAGTCTGTTCTGATTTGGGCTTGCACCTCCAGTCTGGCCCTGCAC	302	
Db	241	CAACTGGCGGGCAACACAGTCTGTTCTGATTTGGGCTTGCACCTCCAGTCTGGCCCTGCAC	300	
QY	303	CGCGCGTCGCAAAATTGTGCGGCGAATTAATCTTCGCGCGGATCAACTGGGTGCCAGCGTG	362	
Db	301	CGCGCGTCGCAAAATTGTGCGGCGAATTAATCTTCGCGCGGATCAACTGGGTGCCAGCGTG	360	
QY	363	GTGTGTGTCGATGGTAGAAACGAAGCGGCGTCCGAAGCTGTAAAGCGCGGTGCACAACTCTT	422	
Db	361	GTGTGTGTCGATGGTAGAAACGAAGCGGCGTCCGAAGCTGTAAAGCGCGGTGCACAACTCTT	420	
QY	423	CTCGCGCAACCGCTCAGTGGGTGATCATTTAACTATCCGCTGGATGACAGGATGCCATT	482	
Db	421	CTCGCGCAACCGCTCAGTGGGTGATCATTTAACTATCCGCTGGATGACAGGATGCCATT	480	

Qy	483	GCTGTGGAAGCTGCGTGCATTAATGTTCTCGCGGTTAATTTCTTGATGTTCTTGACCA	542
Db	481	GCTGTGGAAGCTGCGTGCATTAATGTTCTCGCGGTTAATTTCTTGATGTTCTTGACCA	540
Qy	543	CCCATCAACAGTATTATTTTCTCCCATGAAGACGGTAGCGCACTGGCGGTGGAGCATCTG	602
Db	541	CCCATCAACAGTATTATTTTCTCCCATGAAGACGGTAGCGCACTGGCGGTGGAGCATCTG	600
Qy	603	GTCCGATTGGGTACACGACAAATCGCGTGTTAGCGGCCCATTAAGTTCTGTCTCGCGG	662
Db	601	GTCCGATTGGGTACACGACAAATCGCGTGTTAGCGGCCCATTAAGTTCTGTCTCGCGG	660
Qy	663	CGTCTGCGTCTGGCTGGCTGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG	722
Db	661	CGTCTGCGTCTGGCTGGCTGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG	720
Qy	723	GAACGGGAAGCGCACTGGAGTGCCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGAAT	782
Db	721	GAACGGGAAGCGCACTGGAGTGCCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGAAT	778
Qy	783	GAGG	786
Db	779	GAGG	782

RESULT 5

BQ751220

LOCUS

DEFINITION

ACCESSION

VERSION _____

KEYWORDS

SOURCE

ORGANISM

REFERENCE

REFERENCE

AUTHORS

TITLE

JOURNAL,

COMMENT

111

FEA TYPE

FEATURES

22709

lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN		Query Match		22.9%; Score 755.4; DB 5; Length 806;		Best Local Similarity		97.4%; Pred. No. 6.3e-211;		Matches 790; Conservative 0; Mismatches 16; Indels 5; Gaps 2;	
Qy	3	GACACCATCGAATGTGTGCAAAACCTTTCGCGTATGGCATGATAGCGCCGGAAGAGAGT	62								
Db	1	GACACTTCGAATGGCGCAAAACCTTTCGCGTATGGCATGATAGCGCCGGAAGAGAGT	60								
Qy	63	CAATTACGGGTGGTGAATGTGAAACACAGTAACTGTTATACGATGTCGACAGATATGCCGGT	122								
Db	61	CAATTACGGGTGGTGAATGTGAAACACAGTAACTGTTATACGATGTCGACAGATATGCCGGT	120								
Qy	123	GTCTCTTATCAGACGTTTCGCGGTGTGTAACAGGCGACGACGTTTCTGCGAAAACG	182								
Db	121	GTCTCTTATCAGACGTTTCGCGGTGTGTAACAGGCGACGACGTTTCTGCGAAAACG	180								
Qy	183	CGGGAAAAGTGGAGCGCGATGCGGAGCTGAATTACATTCCTCAACCGGTGCGACAA	242								
Db	181	CGGGAAAAGTGGAGCGCGATGCGGAGCTGAATTACATTCCTCAACCGGTGCGACAA	240								
Qy	243	CAACTGGCGGCAACACAGTCTGTGATGTCGCTGTCACCTCCAGTCTGCGCCCTGCAC	302								
Db	241	CAACTGGCGGCAACACAGTCTGTGATGTCGCTGTCACCTCCAGTCTGCGCCCTGCAC	300								
Qy	303	CGCGCGTCCGAAATGTGCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCGTG	362								
Db	301	CGCGCGTCCGAAATGTGCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCGTG	360								
Qy	363	GTGGTGTGATGATAGACGAGCGGCTGCAAGCCTGTAAGCGCGGTGCACATCTT	422								
Db	361	GTGGTGTGATGATAGACGAGCGGCTGCAAGCCTGTAAGCGCGGTGCACATCTT	420								
Qy	423	CTCGCGCAACGCGTCAGTGGGCTGATCACTAATCACTATCGCTGGATGACAGGATGCCATT	482								
Db	421	CTCGCGCAACGCGTCAGTGGGCTGATCACTAATCACTATCGCTGGATGACAGGATGCCATT	480								
Qy	483	GCTGTGGAAGTGCCTGCACATAATGTTCCGGGTTATTTCTTGATGTCTTGACACAGACA	542								
Db	481	GCTGTGGAAGTGCCTGCACATAATGTTCCGGGTTATTTCTTGATGTCTTGACACAGACA	540								
Qy	543	CCCATCAACAGTATTTTCTCCCATGACGCGTACCGACTGCGGCTGGGAGCATCTG	602								
Db	541	CCCATCAACAGTATTTTCTCCCATGACGCGTACCGACTGCGGCTGGGAGCATCTG	600								
Qy	603	GTCCGATTGGGTACACGCAAAATCGCGTGTAGCGGGCCCAATTAAGTTCTGTCTCGGGG	662								
Db	601	GTCCGATTGGGTACACGCAAAATCGCGTGTAGCGGGCCCAATTAAGTTCTGTCTCGGGG	660								
Qy	663	CGTCTGCGTCTGGCTGGGTGCAATAATCTCACTCGCAATCAAATTCAGCCGATAGCG	722								
Db	661	CGTCTGCGTCTGGCTGGGTGCAATAATCTCACTCGCAATCAAATTCAGCCGATAGCG	720								
Qy	723	GAAACGGGAAGCGACTGAGTGCATATGTCGGGTTTTCACAAACCATGCAATGCTGAAT	782								
Db	721	GAAACGGGAAGCGACTGAGTGCATATGTCGGGTTTTCACAAACCATGCAATGCTGAAT	780								
Qy	783	GAGGCGATCGTTCCACATCGCATGCTGGTTG	813								
Db	778	GAGG---CATGCTTCCATGCGATGCTGGTTG	806								

RESULT 6
CB863541/c
LOCUS HH04A08y HH Hordeum vulgare cDNA clone HH04A08 3-PRIME, mRNA
DEFINITION CB863541 689 bp mRNA linear EST 22-APR-2003
sequence.

CB863541
CB863541.1 GI:30059100
EST.
KEYWORDS
SOURCE
ORGANISM
Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Hordeum.
1 (bases 1 to 689)
Varshney,R.K., Zhang,H., Burton,R., Stein,N., Langridge,P. and
Granel,A.
Barley ESTs from coleoptile tissue
Unpublished (2003)
Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 689 Std Error: 0.00
Plate: 4 row: A column: 8
Seq primer: SP6.
Location/Qualifiers
1...689
/organism="Hordeum vulgare"
/mol_type="mRNA"
/cultivar="Sloop"
/db_xref="GABI:555198"
/db_xref="taxon:4513"
/clone="HH04A08"
/tissue_type="coleoptile"
/dev_stage="coleoptile, 1 day old"
/lab_host="DH10B"
/clone_lib="HH"
/note="Vector: pSPORT; Site 1: SalI (5-end of cDNA);
Site 2: NotI (3-end of cDNA); Due to the cloning system
used blue/white selection for recombinants is not 100 %
reliable. Average insert size is 1.3 Kb."

ORIGIN

Query Match		20.8%; Score 684.8; DB 6; Length 689;		Best Local Similarity		99.6%; Pred. No. 4.2e-190;		Matches 686; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
Qy	21	AAACCTTTTCGGGTATGGCATGATAGCGCCGGAAGAGAGTCAATTGAGGTGGTGAAT	80						
Db	689	AAACCTTTTCGGGTATGGCATGATAGCGCCGGAAGAGAGTCAATTGAGGTGGTGAAT	630						
Qy	81	GTGAAACCAAGTAAAGTTATAGATGTCGACAGTATGCGGTGTTCTTTATCAGACCGTT	140						
Db	629	GTGAAACCAAGTAAAGTTATAGATGTCGACAGTATGCGGTGTTCTTTATCAGACCGTT	570						
Qy	141	TCCCGCGTGGTGAACCGACCGCCAGCTTTCTCGGAAACCGCGGAAAGTGGAAAGCG	200						
Db	569	TCCCGCGTGGTGAACCGACCGCCAGCTTTCTCGGAAACCGCGGAAAGTGGAAAGCG	510						
Qy	201	GGATGGCGGAGCTGAATTACATTCCTCAACCGCTGGGACCAACTGCGCGGCAACAG	260						
Db	509	GGATGGCGGAGCTGAATTACATTCCTCAACCGCTGGGACCAACTGCGCGGCAACAG	450						
Qy	261	TGTTGCTGATTGGCGTTGCCACCTTCCAGTCTGGCCCTGACGCGCGCTGCGAAATTGTC	320						
Db	449	TGTTGCTGATTGGCGTTGCCACCTTCCAGTCTGGCCCTGACGCGCGCTGCGAAATTGTC	390						
Qy	321	GGGCGATTAAATCTCGCGCGGATCAACTGGGTGCCAGGTGGTTCGATGTTAGTA	380						
Db	389	GGGCGATTAAATCTCGCGCGGATCAACTGGGTGCCAGGTGGTTCGATGTTAGTA	330						
Qy	381	CGAGCGCGCTGGAAGCTGTAAAGCGCGGTGACAACTTCTCGCGCAACGCGTCACT	440						
Db	329	CGAGCGCGCTGGAAGCTGTAAAGCGCGGTGACAACTTCTCGCGCAACGCGTCACT	270						

```

QY 441 GGCTGATCAATTAACATATCCGCTGGATGACAGGATGCCATTGCTGTGGAAGCTGCCTGC 500
Db 269 GGCTGATCAATTAACATATCCGCTGGATGACAGGATGCCATTGCTGTGGAAGCTGCCTGC 210
QY 501 ACTAATGTTCCCGCGTATTCTTGTGATGCTCTGACGACACCCCATCAACAGTATTATT 560
Db 209 ACTAATGTTCCCGCGTATTCTTGTGATGCTCTGACGACACCCCATCAACAGTATTATT 150
QY 561 TTCTCCCATGAAGACCGGTACGGACACTGGCGGTGGAGCATCTGCTGCAATTTGGTCAACAG 620
Db 149 TTCTCCCATGAAGACCGGTACGGACACTGGCGGTGGAGCATCTGCTGCAATTTGGTCAACAG 90
QY 621 CAAATCGCGCTGTAGCGGCGCCATTAACTTCTGCTCGGCGCGTCTGCTGCTGCTGCTG 680
Db 89 CAAATCGCGCTGTAGCGGCGCCATTAACTTCTGCTCGGCGCGTCTGCTGCTGCTGCTGCTG 30
QY 681 TGGCATAAATATCTCACTCGCAATCAAT 709
Db 29 TGGCATAAATATCTCACTCGCAATCAAT 1

RESULT 7
AL038548/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL038548 712 bp mRNA linear EST 04-SEP-2003
DKFZ566E0346_r1 566 (synonym: hfkd2) Homo sapiens cDNA clone
DKFZ566E0346 5', mRNA sequence.
AL038548
AL038548.1 GI:5407738
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Ottenwaelder,B., Obermaier,B., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
EST (Ottenwaelder, et al.)
Unpublished (1999)
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by MedGenomix (Martinried/Germany) within the cDNA
sequencing consortium of the German Genome Project. No sl sequence
available.
This clone (DKFZ566E0346) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source
1..712
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZ566E0346"
/tissue_type="kidney"
/dev_stage="fetal"
/lab_host="X1-2blue"
/clone_lib="566 (synonym: hfkd2)"
/note="Vector: pAMP1; Site_1: NotI; Site_2: SalI"

Query Match 20.5%; Score 676.4; DB 1; Length 712;
Best Local Similarity 99.6%; Pred. No. 1.3e-187;
Matches 699; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
QY 33 GGTATGCGATAGCGCCCGGAGAGAGTCAATTGAGGTGGTGAATGGAACAGTA 92
Db 712 GGTATGCGATAGCGCCCGGAGAGAGTCAATTGAGGTGGTGAATGGAACAGTA 653
QY 93 ACCTTATACATGTCGAGAGTATGCCGGTGTCTTATCAGACCGTTCCCGCGGTG 152

```

```

Db 652 AGTTATACAGTATGCGCAGAGTATGCCGGTGTCTCTTATACAGCCGTTTCCCGGTGGT 593
QY 153 AACCAGGCCACCGCTTTCTGCGAAAACCGCGG-AAAAAGTGAAGCGCGGATGGCGGA 211
Db 592 AACCAGGCCACCGAGTTTCTGCGAAAACCGCGGAAAAAGTGAAGCGCGGATGGCGGA 533
QY 212 GCTGAATTAACATTTCCCAACCGCGTGGGCACAACAACCTGGCGGGCAACAGTCTGTTCTGAT 271
Db 532 GCTGAATTTACATTTCCCAACCGCGTGGGCACAACAACCTGGCGGGCAACAGTCTGTTCTGAT 473
QY 272 TGGGTTTGGCCACCTCCAGTCTGGCCCTGCGAC-GGCGCTGCGCAAAATTTGTCGGCGGATTA 330
Db 472 TGGGTTTGGCCACCTCCAGTCTGGCCCTGCGACGGCGCCCTGCGCAAAATTTGTCGGCGGATTA 413
QY 331 AATCTCGCGCGGATCAACTGGGTGCCAGCGTGGTGTGTCGATGCTAGAAACGAAGCGCGC 390
Db 412 AATCTCGCGCGGATCAACTGGGTGCCAGCGTGGTGTGTCGATGCTAGAAACGAAGCGCGC 353
QY 391 TCGAAGCGCTGTAAAGCGCGGTGCACAATCTTCTCGCGCAACCGGTGAGTCA 450
Db 352 TCGAAGCGCTGTAAAGCGCGGTGCACAATCTTCTCGCGCAACCGGTGAGTCA 293
QY 451 TTAACATATCGCTGGATGACGAGATGCCATTGCTGTGGAAGCTGCTGCACTAAATGTTTC 510
Db 292 TTAACATATCGCTGGATGACGAGATGCCATTGCTGTGGAAGCTGCTGCACTAAATGTTTC 233
QY 511 CGCGTTATTCTTGTGATGCTCTGACGACACCCATCAACAGTATTATTTCTCCCATG 570
Db 232 CGCGTTATTCTTGTGATGCTCTGACGACACCCATCAACAGTATTATTTCTCCCATG 173
QY 571 AAGACGCTATCGCACTGGGCGTGGAGCATCTGTCGCATTGGTCCACCAATTCGCGC 630
Db 172 AAGACGCTATCGCACTGGGCGTGGAGCATCTGTCGCATTGGTCCACCAATTCGCGC 113
QY 631 TGTAGCGGCGCCATTAAAGTTCTGTCGCGCGCTCTGGCTCTGGCTGGCATAAAT 690
Db 112 TGTAGCGGCGCCATTAAAGTTCTGTCGCGCGCTCTGGCTCTGGCTGGCATAAAT 53
QY 691 ATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAG 732
Db 52 ATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAG 11

AV080106 710 bp DNA linear GSS 06-NOV-2002
AY080106 Scripps Pier (La Jolla, CA) uncultured virus community
uncultured marine virus genomic clone SIO51p3G6L, genomic survey
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

AV080106.1 GI:24745302
uncultured marine virus
Viruses; environmental samples.
1 (bases 1 to 710)
Breitbart,M., Salamon,P., Andresen,B., Mahaffy,J.M., Segall,A.M.,
Mead,D., Azam,F. and Rohwer,F.
Genomic analysis of uncultured marine viral communities
Proc. Natl. Acad. Sci. U.S.A. 99 (22), 14250-14255 (2002)
22294988
Contact: Rohwer F
Biology Dept.
San Diego State University
5500 Campanile Dr, San Diego, CA 92102, USA
Tel: 6195941336
Fax: 619595676
Email: forest@sunstroke.sdsu.edu
Class: shotgun.
Location/Qualifiers
1..710
/organism="uncultured marine virus"

```

```
/mol_type="genomic DNA"
/db_xref="taxon:18617"
/clone="SI051pJG6L"
/clone_lib="Scripps Pier (La Jolla, CA) uncultured virus
community"
/note="Marine viruses were isolated from 200 liters of
surface seawater using a combination of differential
filtration and density-dependent gradient centrifugation.
Linker-amplified shotgun libraries were constructed by
randomly shearing the total marine viral community DNA,
end-repairing, ligating dsDNA linkers to the ends, and
amplifying the fragments using Vent DNA polymerase. The
resulting fragments were ligated into the pSMART vector
and electroporated into MC12 cells (Lucigen; Middleton,
WI)."
```

ORIGIN

```
Query Match      20.5%; Score 676; DB 8; Length 710;
Best Local Similarity 98.9%; Pred. No. 1.7e-187;
Matches 690; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

QY 400 GTAAAGCGCGTGACAACTCTTCGCGCAACGGCTCAGTGGGCTGATCACTAACTATC 459
DB |||
DB 8 GTGCTGACGCGTGACAACTCTTCGCGCAACGGCTCAGTGGGCTGATCACTAACTATC 67
QY 460 CGCTGGATGACAGATGCGATCTGTGTGGAAAGCTGCGTGCACCTAATGTTCCGGGTTAT 519
DB |||
DB 68 CGCTGGATGACAGATGCGATCTGTGTGGAAAGCTGCGTGCACCTAATGTTCCGGGTTAT 127
QY 520 TTCTTTGATGTCCTGACACGACACCATCAACAGATATTATTTCTCCCATGAAGACGGTA 579
DB |||
DB 128 TTCTTTGATGTCCTGACACGACACCATCAACAGATATTATTTCTCCCATGAAGACGGTA 187
QY 580 CGCGACTGGCGTGAGCATCTGTGCGCATCTGGTCCAGCAAAATCGCGCTGTTAGCGG 639
DB |||
DB 188 CGCGACTGGCGTGAGCATCTGTGCGCATCTGGTCCAGCAAAATCGCGCTGTTAGCGG 247
QY 640 GCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTC 699
DB |||
DB 248 GCCCATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTC 307
QY 700 GCAATCAATTCAGCCGATAGCGGAACGGAAGGCGACTGGAGTGCCATGTCGGGTTTC 759
DB |||
DB 308 GCAATCAATTCAGCCGATAGCGGAACGGAAGGCGACTGGAGTGCCATGTCGGGTTTC 367
QY 760 AACAAACCATGCAAAATGCTGAATGAGGCGATCGTTCCCACTCGCATGCTGGTTGCCAACG 819
DB |||
DB 368 AACAAACCATGCAAAATGCTGAATGAGGCGATCGTTCCCACTCGCATGCTGGTTGCCAACG 427
QY 820 ATCAGATGCGCTGGGCGCAATGCGGCGCATTTACGAGTCCGGGCTGCGGCTGGTGGCGG 879
DB |||
DB 428 ATCAGATGCGCTGGGCGCAATGCGGCGCATTTACGAGTCCGGGCTGCGGCTGGTGGCGG 487
QY 880 ATATCTCGGTAGTGAGTACGAGTACCGAAGACAGCTCATGTTATATCCCGCGTTAA 939
DB |||
DB 488 ATATCTCGGTAGTGAGTACGAGTACCGAAGACAGCTCATGTTATATCCCGCGTTAA 547
QY 940 CCACCATCAAAAGGATTTTCGCTGCTGGGCGCAACAGCGTGGACCGCTTGTGTCACAC 999
DB |||
DB 548 CCACCATCAAAAGGATTTTCGCTGCTGGGCGCAACAGCGTGGACCGCTTGTGTCACAC 607
QY 1000 TCTCTCAGGGCGAGCGGTGAAGGGCAATCAAGCTGTGTCCTGCTCACTGTTGAAAGAA 1059
DB |||
DB 608 TCTCTCAGGGCGAGCGGTGAAGGGCAATCAAGCTGTGTCCTGCTCACTGTTGAAAGAA-AA 666
QY 1060 AAACCACTCTGGGCGCCCAATAGCAACCGCTCTCTCC 1097
DB |||
DB 667 AAACCACTCTGGGCGCCCAATAGCAACCGCTCTCTCC 704
```

RESULT 9

AL044483/c

LOCUS

660 bp mRNA linear EST 04-SEP-2003

```
DEFINITION
DKFZp4341102_s1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION
DKFZp4341102.3, mRNA sequence.
VERSION
AL044483.1 GI:5432701
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 660)
AUTHORS
Ansorge,W., Benes,V., Krieger,S., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
TITLE
EST (Ansorge, Benes, et al.)
JOURNAL
Unpublished (1999)
COMMENT
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 3' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No r1 sequence available.
This clone (DKFZp4341102) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..660
/morganism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp4341102"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
```

FEATURES

source

ORIGIN

```
Query Match      19.6%; Score 647.4; DB 1; Length 660;
Best Local Similarity 99.7%; Pred. No. 4.5e-179;
Matches 659; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 85 AACCAAGTAACTATACGATGTCCAGAGATGCGCGGTGCTCTTTATCAGACCGTTTCCC 144
DB |||
DB 660 AACCAAGTAACTATACGATGTCCAGAGATGCGCGGTGCTCTTTATCAGACCGTTTCCC 601
QY 145 GGTGTTGACACGAGCCAGCCACGTTTCTGCGAAACCGGGGAAAGTGGAGCGGCGA 204
DB |||
DB 600 GGTGTTGACACGAGCCAGCCACGTTTCTGCGAAACCGGGGAAAGTGGAGCGGCGA 541
QY 205 TGGCGGAGCTGAATATACATTTCCCAACCGCTGGGCAACCACTGGCGGCAACAGTCGT 264
DB |||
DB 540 TGGCGGAGCTGAATATACATTTCCCAACCGCTGGGCAACCACTGGCGGCAACAGTCGT 481
QY 265 TGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCACGCGCGCTCGCAAAATTTGCGCG 324
DB |||
DB 480 TGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCACGCGCGCTCGCAAAATTTGCGCG 421
QY 325 CGATTAAATCTCGCGCGGATCACTGGGTGCGAGCGTGGTGTGTCGATGGTAGAACGAA 384
DB |||
DB 420 CGATTAAATCTCGCGCGGATCACTGGGTGCGAGCGTGGTGTGTCGATGGTAGAACGAA 361
QY 385 GGGCGGTGGAACGCTGTAAGCGGCGGTGCACAAATCTTCTCGCGCAACCGCTCAGTGGGC 444
DB |||
DB 360 GGGCGGTGGAACGCTGTAAGCGGCGGTGCACAAATCTTCTCGCGCAACCGGTGAGTGGGC 301
QY 445 TGATCATTAATCATCCGCTGGATGACCAAGATGCCATTGCTGTGGAAGCTGCTGCACATA 504
DB |||
DB 300 TGATCATTAATCATCCGCTGGATGACCAAGATGCCATTGCTGTGGAAGCTGCTGCACATA 241
QY 505 ATGTTCCGGCGTTATTTCTTGTGATGCTCTGACACGACACCCCATCAACAGTATATTTTCT 564
```



```

Db      240  ATGTTCCGGGTTATTTCTTTGATGTTCTCTGACCAGACCCATCAACAGTATATTTTCT 181
QY      565  CCCATGAAGACGGTAGCGACTGGCGGTGGAGCATCTGTCGATTTGGGTACACAGAAA 624
Db      180  CCATGAAGACGGTAGCGACTGGCGGTGGAGCATCTGTCGATTTGGGTACACAGAAA 121
QY      625  TC CGCGTGTAGCGGCCCCATTAAGTTCTGTCCTCGCGCGGTCTGGGTCTGGCTGGCTGGC 684
Db      120  TC CGCGTGTAGCGGCCCCATTAAGTTCTGTCGCG- GCGTCTGGCTGGCTGGCTGGC 62
QY      685  ATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCGAAACGGGAAGCGACTGGAGTG 744
Db      61  ATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCGAAACGGGAAGCGACTGGAGTG 2
QY      745  C 745
Db      1  C 1

RESULT 10
BJ063872/c
LOCUS   BJ063872 NTBB Mochii normalized Xenopus tailbud library Xenopus
DEFINITION
ACCESSION BJ063872
VERSION   BJ063872.1 GI:17471074
KEYWORDS EST.
SOURCE   Xenopus laevis (African clawed frog)
ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
    Xenopodinae; Xenopus; Xenopus.
REFERENCE
    1 (bases 1 to 641)
    Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and
    Kohara,Y.
    Expressed genes in X. laevis embryo
    Unpublished (2001)
    Contact: Tadasu Shin-i
    Center For Genetic Resource Information
    National Institute of Genetics
    1111 Yata, Mishima, Shizuoka 411-8540, Japan
    Tel: 81-559-81-6856
    Fax: 81-559-81-6855
    Email: tsunigenes.nig.ac.jp
    The information of this clone is available through the following
    URL.
    http://xenopus.nibb.ac.jp.
FEATURES             source
    Location/Qualifiers
        1..641
            /organism="Xenopus laevis"
            /mol_type="mRNA"
            /db_xref="taxon:8355"
            /clone="XL077K02"
            /tissue_type="whole embryo"
            /dev_stage="stage 25"
            /clone_lib="NIBB Mochii normalized Xenopus tailbud
            library"
ORIGIN
Query Match      19.4%; Score 641; DB 4; Length 641;
Best Local Similarity 100.0%; Pred. No. 3.5e-177;
Matches 641; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      30  CGCGGTATGGCATGATAGCGCCCGAAGAGAGTCAATTCAGGGTGGTGAATGTGAACCA 89
Db      641  CGCGGTATGGCATGATAGCGCCCGAAGAGAGTCAATTCAGGGTGGTGAATGTGAACCA 582
QY      90  GTAACGTTATACGATGTCGAGAGTATGCCGGTGTCTCTTTATCAGACCGTTTCCCGGTG 149
Db      581  GTAACGTTATACGATGTCGAGAGTATGCCGGTGTCTCTTTATCAGACCGTTTCCCGGTG 522
QY      150  GTGAACCGAGCCACCGTCTTCTCGCAAAACCGGGGAAAAAGTGGGAAGCGCGGTGGCG 209

```

```

Db      521  GTGAACCGAGCGAGCCAGTTCCTGCGAAAAACGGGGAAAAAGTGAAGCGCGATGCGC 462
QY      210  GAGCTGAATTACATTTCCCAACCGGTGSCACAACTGGCGGGCAACACAGTCGTTGCTG 269
Db      461  GAGCTGAATTACATTTCCCAACCGGTGSCACAACTGGCGGGCAACACAGTCGTTGCTG 402
QY      270  ATTGGCGTTGCCACTCCAGTCTGGCCCTGACGCGCGCTGCGCAAAATTTGTCGCGCGCAT 329
Db      401  ATTGGCGTTGCCACTCCAGTCTGGCCCTGACGCGCGCTGCGCAAAATTTGTCGCGCGCAT 342
QY      330  AAATCTCGCGCGCATCAACTGGGTGCCAGCGGTGCTGTCGATGGTAGAACGAAGCGGC 389
Db      341  AAATCTCGCGCGCATCAACTGGGTGCCAGCGGTGCTGTCGATGGTAGAACGAAGCGGC 282
QY      390  GTCGAAGCCTGTAAAGCGCGGTGCACAAATCTTCTCGCGCAACCGCTCAGTGGGCTGATC 449
Db      281  GTCGAAGCCTGTAAAGCGCGGTGCACAAATCTTCTCGCGCAACCGCTCAGTGGGCTGATC 222
QY      450  ATTAACCTATCGCTGGATGACAGGATGCCATGCTGTGGAAGCTGCTGCACCTAAATGTT 509
Db      221  ATTAACCTATCGCTGGATGACAGGATGCCATGCTGTGGAAGCTGCTGCACCTAAATGTT 162
QY      510  CCGCGCTTATTTCTTGATGTCCTGACCGACACCCCATCAACAGTATATTTCTCCCAT 569
Db      161  CCGCGCTTATTTCTTGATGTCCTGACCGACACCCCATCAACAGTATATTTCTCCCAT 102
QY      570  GAAGACGGTACGGCACTGGCGGTGGAGCATCTGTCGCAATGGGTCAACGACAAATCGCG 629
Db      101  GAAGACGGTACGGCACTGGCGGTGGAGCATCTGTCGCAATGGGTCAACGACAAATCGCG 42
QY      630  CTGTTAGCGGGCCCATTAAGTTCTGTCGCGCGCTGCGG 670
Db      41  CTGTTAGCGGGCCCATTAAGTTCTGTCGCGCGCTGCGG 1

RESULT 11
CB863814/c
LOCUS   CB863814 HH Hordeum vulgare cDNA clone HH07B20 3-PRIME, mRNA
DEFINITION
ACCESSION CB863814
VERSION   CB863814.1 GI:30058373
KEYWORDS EST.
SOURCE   Hordeum vulgare
ORGANISM
    Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
    Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
    Poideae; Triticeae; Hordeum.
    Varshney,R.K., Zhang,H., Burton,R., Stein,N., Langridge,P. and
    Graner,A.
    Barley ESTs from coleoptile tissue
    Unpublished (2003)
    Contact: Stein Nils
    Molecular Markers Group, Department Genbank
    Institute of Plant Genetics and Crop Plant Research (IPK)
    Corrensstr. 3, 06466, Gatersleben, Germany
    Tel: 039482-5522
    Fax: 039482-5595
    Email: stein@ipk-gatersleben.de
    Insert Length: 624 Std Error: 0.00
    Plate: 7 row: B column: 20
    Seq primer: SP6.
FEATURES             source
    Location/Qualifiers
        1..624
            /organism="Hordeum vulgare"
            /mol_type="mRNA"
            /culturivar="Sloop"
            /db_xref="GABI:555471"
            /db_xref="taxon:4513"
            /clone="HH07B20"
            /tissue_type="coleoptile"

```


EST (Duesterhoeft, et al.)		Unpublished (1999)		Contact: MIPS		Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany		This is the 3' sequence of the clone insert		Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); sequenced by Qiagen within the cDNA sequencing consortium of the German Genome Project. r1 sequence also available.		This clone is available at the RZPD in Berlin.		Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.											
FEATURES		source		1..794		/organism="Homo sapiens"		/mol_type="mRNA"		/db_xref="taxon:9606"		/clone="DKFZp434B075"		/tissue_type="testis"		/dev_stage="adult"		/lab_host="DH10B"		/clone_lib="434 (synonym: htes3)"		/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"			
ORIGIN		Query Match		Best Local Similarity		18.5%; Score 612; DB 1; Length 794;		Matches 634; Conservative 0; Mismatches 0; Indels 2; Gaps 2;		QY		569		TGAAGACGGTAGCGACTGGGCGTGGAGCATCTGTCGATTTGGTCCACCAATATCGC		628		Db		794		TGAAGACGGTAGCGACTGGGCGTGGAG-ATCTGTCGCAATTTGGGTCCACCAATATCGC		736	
QY		629		GCTGTTACGCGGCCCATTAAGTTCTGTCGCGCGGTCTCGCTCTGGCTGGCTGGCATAA		688		Db		735		GCTGTTACGCGGCCCATTAAGTTCTGTCGCGCGGTCTG-CTCTGGCTGGCTGGCATAA		677		QY		689		ATATCTCACTCGCAATCAAAATTCAGCCGATAGCGAAGCGGAAAGCGACTGGAGTGCCAT		748			
Db		676		ATATCTCACTCGCAATCAAAATTCAGCCGATAGCGAAGCGGAAAGCGACTGGAGTGCCAT		617		QY		749		GTCGGGTTTCAACAAACCATGCAATGCTGATGAGGGCATCGTTCACACTGCGATGCT		808		Db		616		GTCGGGTTTCAACAAACCATGCAATGCTGATGAGGGCATCGTTCACACTGCGATGCT		557			
QY		809		GTTTCCCAACGATCAGATGGCGCTGGCGCAATGCGCGCATTTACCGAGTCCGGGCTGCG		868		Db		556		GTTTCCCAACGATCAGATGGCGCTGGCGCAATGCGCGCATTTACCGAGTCCGGGCTGCG		497		QY		869		CGTTGGTGGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATAT		928			
Db		496		CGTTGGTGGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATAT		437		QY		929		CCGCGCGTTAACCCACCATCAAAACAGGATTTTCGCTGCTGGGCGAAGCCAGGCTGGACCG		988		Db		436		CCGCGCGTTAACCCACCATCAAAACAGGATTTTCGCTGCTGGGCGAAGCCAGGCTGGACCG		377			
QY		989		CTTGTGCAACTCTCTCAGGGCGAGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACT		1048		Db		376		CTTGTGCAACTCTCTCAGGGCGAGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACT		317		QY		1049		GGTGAAGAAAGAAACACACCTGGCGCCCAATAACGCAAAACCGCTCTCCCGCGCGTGGC		1108			
Db		316		GGTGAAGAAAGAAACACACCTGGCGCCCAATAACGCAAAACCGCTCTCCCGCGCGTGGC		257		QY		1109		CGATTCAATTAATGACGCTGGCAACAGATTTTCCGCACTGGAAAGCGGCGAGTAGCGCA		1168		Db		256		CGATTCAATTAATGACGCTGGCAACAGATTTTCCGCACTGGAAAGCGGCGAGTAGCGCA		197			
QY		1169		ACGCAATTAATGAGTTAGTCTCACTCATATTAGGCAC		1204		Db		196		ACGCAATTAATGAGTTAGTCTCACTCATATTAGGCAC		161											

```

RESULT 13
CL655575/c
LOCUS
DEFINITION
    CL655575
    PRI0124a.C01 - PRI0124a.B21 (752) Mixed stage fosmid library of P.
    pacificus var. California Pristionchus pacificus genomic, genomic
    survey sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
    GSS.
    CL655575.1 GI:50134897
    Pristionchus pacificus
    Pristionchus pacificus
    Pristionchus pacificus
    Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
    Neodiplogasteridae; Pristionchus.
REFERENCE
    1 (bases 1 to 752)
    Srinivasan, J., Otto, G.W., Kahlow, U., Geisler, R. and Sommer, R.J.
    AppADB: an AcedB database for the nematode satellite organism
    Pristionchus pacificus
    Nucleic Acids Res. 32 (1), D421-D422 (2004)
    Contact: Sommer RJ
    Evolutionary Biology
    Max-Planck-Institute for Developmental Biology
    Spemannstr. 37-39, Tuebingen D-72076, Germany
    Tel: 00497071601371
    Fax: 00497071601498
    Email: raif.sommer@tuebingen.mpg.de
    This library was generated at Caltech, Pasadena, USA and end
    sequenced at Vancouver, Canada.
    Seq primer: T7
    Class: fosmid ends.
FEATURES
    source
        Location/Qualifiers
        1..752
        /organism="Pristionchus pacificus"
        /mol_type="genomic DNA"
        /strain="California"
        /db_xref="taxon:54126"
        /clone_lib="Mixed stage fosmid library of P. pacificus
        var. California"
        /note="Vector: pBpifos-5 Fosmid vector"
ORIGIN
    Query Match 18.3%; Score 603; DB 9; Length 752;
    Best Local Similarity 98.9%; Pred. No. 6e-166;
    Matches 628; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

Qy 1996 TTCACCTGGCCGCTGATTCGCTGAC -GGGGGTTATGCGTTCAAGTATGAAACGGCAA 2054
Db 752 TTCACCTGGCCGCTGATTCGCTGAC -GGGGGTTATGCGTTCAAGTATGAAACGGCAA 693

Qy 2055 GTACGACATTAAAGACGTTGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCT 2114
Db 692 GTACGACATTAAAGACGTTGGCGTGGATAACGCTGGCGCGAAAGCGGGTCTGACCTTCCT 633

Qy 2115 GGTTGACCTGATTAATAAACACATCATGATGCGAGACACCGATTACTCCATCGAGAAGC 2174
Db 632 GGTTGACCTGATTAATAAACACACATCATGATGCGAGACACCGATTACTCCATCGAGAAGC 573

Qy 2175 TGCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCGCGTGGCGATGTTCCACAT 2234
Db 572 TGCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCGCGTGGCGATGTTCCACAT 513

Qy 2235 CGACACAGCAAGTGAATATGTTGTTAAACGTTACTGCCGACCTTCAAGGTTCAACCATC 2294
Db 512 CGACACAGCAAGTGAATATGTTGTTAAACGTTACTGCCGACCTTCAAGGTTCAACCATC 453

Qy 2295 CAACCCGTTGTTGGCGTGTGAGCGAGGTATTAACGCCCGCAGTCCGACAAAGGCT 2354
Db 452 CAACCCGTTGTTGGCGTGTGAGCGAGGTATTAACGCCCGCAGTCCGACAAAGGCT 393

Qy 2355 GGCAAAAGAGTCTCTCGAAACATATCTGCTGACTGATGAAGGTTCTGGAACGGGTTAA -TA 2413
Db 392 GGCAAAAGAGTCTCTCGAAACATATCTGCTGACTGATGAAGGTTCTGGAACGGGTTAA 333

```

```

Qy 2414 AAGACAAACCGCTGGGTGCGGTAGCTTACGAGGAAGAGTTGCGGAAAGATC 2473
Db 332 AGGCCAACCCGCTGGGTGCGGTAGCTTACGAGGAAGAGTTGCGGAAAGATC 273

Qy 2474 CACGTATTGCGCGCCACCATGGAAACGCCCAAGAGTGAATCATGCCGAACATCCGCG 2533
Db 272 CACGTATTGCGCGCCACCATGGAAACGCCCAAGAGTGAATCATGCCGAACATCCGCG 213

Qy 2534 AGATGTCGCGTTTCTGTTATGCGGTACTGCGGTGATCAACGCCCGCCAGCGGTGTC 2593
Db 212 AGATGTCGCGTTTCTGTTATGCGGTACTGCGGTGATCAACGCCCGCCAGCGGTGTC 153

Qy 2594 AGACTGTCGATGAAGCCCTGAAGACGCGCAGACT 2628
Db 152 AGACTGTCGATGAAGCCCTGAAGACGCGCAGACT 118

RESULT 14
AV080094/c
LOCUS
DEFINITION
    AV080094
    Scripps Pier (La Jolla, CA) uncultured virus community
    uncultured marine virus genomic clone SIO51p3D4L, genomic survey
    sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
    GSS.
    AV080094.1 GI:24745276
    uncultured marine virus
    uncultured marine virus
    Viruses; environmental samples.
REFERENCE
    1 (bases 1 to 608)
    Breitbart, M., Salamon, P., Andresen, B., Mahaffy, J.M., Segall, A.M.,
    Mead, D., Azam, F. and Rohwer, F.
    Genomic analysis of uncultured marine viral communities
    Proc. Natl. Acad. Sci. U.S.A. 99 (22), 14250-14255 (2002)
    22294988
    12384570
    Contact: Rohwer F
    Biology Dept.
    San Diego State University
    5500 Campanile Dr, San Diego, CA 92102, USA
    Tel: 6195941336
    Fax: 619595676
    Email: forest@sunstroke.sdsu.edu
    Class: shotgun.
FEATURES
    source
        Location/Qualifiers
        1..608
        /organism="uncultured marine virus"
        /mol_type="genomic DNA"
        /db_xref="taxon:186617"
        /clone="SIO51p3D4L"
        /clone_lib="Scripps Pier (La Jolla, CA) uncultured virus
        community"
        /note="Marine viruses were isolated from 200 liters of
        surface seawater using a combination of differential
        filtration and density-dependent gradient centrifugation.
        Linker-amplified shotgun libraries were constructed by
        randomly shearing the total marine viral community DNA,
        end-repairing, ligating dEDNA linkers to the ends, and
        amplifying the fragments using Vent DNA polymerase. The
        resulting fragments were ligated into the pSMART vector
        and electroporated into MC12 cells (Lucigen; Middleton,
        WI)"
ORIGIN
    Query Match 17.9%; Score 589.8; DB 8; Length 608;
    Best Local Similarity 99.7%; Pred. No. 4.5e-162;
    Matches 591; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 287 CAGTCTGGCCCTGCACGCCCATCGCAAAATTGTCGGCGGATTAAATCTCGCGCCGATCA 346
Db 608 CAGTCTGGCCCTGCACGCCCATCGCAAAATTGTCGGCGGATTAAATCTCGCGCCGATCA 549

```

347	QY	ACTGGGTGCCACGGTGGTGGTGCATGGTATAGAACGAAGCGCGTCCGAGCCTGTAAAGC	406
548	Db	ACTGGGTGCCACGGTGGTGGTGCATGGTATAGAACGAAGCGCGTCCGAGCCTGTAAAGC	489
407	QY	GGCGGTGCACAATCTTCTCGGCGAACGGCTCAGTGGGCTGATCATTAACATATCCGCTGGA	466
488	Db	GGCGGTGCACAATCTTCTCGGCGAACGGCTCAGTGGGCTGATCATTAACATATCCGCTGGA	429
467	QY	TGACCAAGGATGCCATTTGCTGTGGAAAGCTGCGCTGCACCTAATGTTCCGGCGGTTATTTCTTGA	526
428	Db	TGACCAAGGATGCCATTTGCTGTGGAAAGCTGCGCTGCACCTAATGTTCCGGCGGTTATTTCTTGA	369
527	QY	TGTCTCTGACCAAGACCCATCAACAGATTAATATTTCTCCCATGAAGACGGTACGGCACT	586
368	Db	TGTCTCTGACCAAGACCCATCAACAGATTAATATTTCTCCCATGAAGACGGTACGGCACT	309
587	QY	GGCGGTGGAGCATCTGTGTGCGATTTGGGTCAACAGCAAAATCGCGCTGTGTAGCGGGCCCATTT	646
308	Db	GGCGGTGGAGCATCTGTGTGCGATTTGGGTCAACAGCAAAATCGCGCTGTGTAGCGGGCCCATTT	249
647	QY	AAGTTTCTGTCGCGCGCTGTCGCTCGCTGGCATAAATATCTCACTCGCAATCA	706
248	Db	AAGTTTCTGTCGCGCGCTGTCGCTGGCATAAATATCTCACTCGCAATCA	189
707	QY	AATTACGCCGATAGCGGAAACGGGAAGGCGCATGGAGTGCCATGTCGGTTTTCAACAAAC	766
188	Db	AATTACGCCGATAGCGGAAACGGGAAGGCGCATGGAGTGCCATGTCGGTTTTCAACAAAC	129
767	QY	CATGCAAAATGCTGAATGAGGCGATCGTTCCCACTCGCATGCTGGTTCGCAACCATCAGAT	826
128	Db	CATGCAAAATGCTGAATGAGGCGATCGTTCCCACTCGCATGCTGGTTCGCAACCATCAGAT	69
827	QY	GGCGCTGGCGCAATTCGCGGCCATTACCGATGTCGGGCTGCGGCTTGGTCGG	879
68	Db	GGCGCTGGCGCAATTCGCGGCCATTACCGATGTCGGGCTGCGGCTTGGTCGG	16

RESULT 15	AL037742	589 bp	mRNA	linear	EST 04-SEP-2000
AL037742/c	DKFp564C157.r1.564	(synonym: hfbr2)	Homo sapiens	cdNA clone	
LOCUS	DKFp564C157.5'				
DEFINITION	mRNA sequence.				
ACCESSION	AL037742				
VERSION	AL037742.1	GI:5407088			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Metazoa; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 589)				
AUTHORS	Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and Wiemann,S.				
TITLE	EST (Bloecker, et al.)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: MIPS				
	MIPS				

```

FEATURES
source
1. .589
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKP20564C157"
Location/Qualifiers
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
This clone (DKP20564C157) is available at the RZPD in Berlin.
No 61 sequence available.
German Genome Project.
Braunschweig/Germany) within the cDNA sequencing consortium of the
sequenced by GBF (National Research Centre for Biotechnology Ltd.
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
This is the 5' sequence of the clone insert
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany

```

```
/tissue_type="brain"  
/dev_stage="fetal"  
/lab_host="X1-2blue"  
/clone_lib="564 (synonym: hibr2)"  
/note="vector: DMP1; Site 1: NotI; Site 2: SalI"
```

ORIGIN

Query Match 17.8%; Score 586.4; DB 1; Length 589;

Best Local Similarity	99.7%;	Pred. No. 4.5e-161;	Length 507
Matches	587;	Conservative	0;
Mismatches	2;	Indels	0;
Gaps	0		

Qy	517	TATTTCTTGATGCTCTGACGAGACCCCATCAACAGTATTATTTCTCCCATGAGACG	576
Db	589	TATTTCTTGATGCTCTGACGAGACCCCATCAACAGTATTATTTCTCCCTTGAAGACG	530
Qy	577	GTACGACATGGGGCTGGAGCATCTGGTCGATTTGGGTACACAGCAAAATCGCGCTGTTAG	636
Db	529	GTACGACATGGGGCTGGAGCATCTGGTCGATTTGGGTACACAGCAAAATCGCGCTGTTAG	470
Qy	637	CGGGCCCAATTAAGTTCTGTCGCGGGCTCTGCGTCTGGCTGGGTGCGATAAATATCTCA	696
Db	469	CGGGCCCAATTAAGTTCTGTCGCGCGCTCTGCGTCTGGCTGGGTGCGATAAATATCTCA	410
Qy	697	CTCGCAATCAAAATTCAGCCGATACCGGAAACGGGAAGGCGACTGAGTCCCATGTCCGGTT	756
Db	409	CTCGCAATCAAAATTCAGCCGATACCGGAAACGGGAAGGCGACTGAGTCCCATGTCCGGTT	350
Qy	757	TTCAACAAACCATGCAAAATGCTGAATCAGGGCATCGTTCCACATGCGATGCTGGTTGCCA	816
Db	349	TTCAACAAACCATGCAAAATGCTGAATCAGGGCATCGTTCCACATGCGATGCTGGTTGCCA	290
Qy	817	ACGATCAGATGGCGCTGGGGCGCAATGGCGGCATTACCGAGTCCGGGCTGCGCGTTGGTG	876
Db	289	ACGATCAGATGGCGCTGGGGCGCAATGGCGGCATTACCGAGTCCGGGCTGCGCGTTGGTG	230
Qy	877	CGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCGT	936
Db	229	CGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCGT	170
Qy	937	TACCCACCATCAACAGGATTTTCGCTGCTGGGGCAAAACAGCGTGGAACCGCTGTGTCG	996
Db	169	TACCCACCATCAACAGGATTTTCGCTGCTGGGGCAAAACAGCGTGGAACCGCTGTGTCG	110
Qy	997	AACTCTCTCAGGGCCAGCGGTGAAGGCAATCAGCTGTGTCCCGCTCTCACTGGTGA AAA	1056
Db	109	AACTCTCTCAGGGCCAGCGGTGAAGGCAATCAGCTGTGTCCCGCTCTCACTGGTGA AAA	50
Qy	1057	GA AAAACCAACCTGGCGCCCAATAACGAAACCGGCTCTCCCGCGCGGCTT	1105
Db	49	GA AAAACCAACCTGGCGCCCAATAACGAAACCGGCTCTCCCGCGCGGCTT	

Search completed: August 3, 2005, 02:02:50
Job time : 6745.2 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 1309.2 Seconds
(without alignments)
16314.373 Million cell updates/sec

Title: US-09-765-555B-16
Perfect score: 3300
Sequence: 1 ccgacaccatgaatggtgc.....acgagcttcggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/FCI_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3300	100.0	3300	10	US-09-765-555-16
2	3216.8	97.5	3300	10	US-09-765-555-18
3	3207.2	97.2	3300	10	US-09-765-555-14
4	3207.2	97.2	3300	10	US-09-765-555-15
5	3207.2	97.2	3300	10	US-09-765-555-17
6	2705.8	82.0	7373	18	US-10-257-384A-1
7	2705.8	82.0	8101	17	US-10-149-472-5

8	2701	81.8	6648	19	US-10-343-859-8	Sequence 8, Appli
9	2698	81.8	5191	19	US-10-343-859-9	Sequence 9, Appli
10	2690.4	81.5	7112	18	US-10-263-153-40	Sequence 40, Appl
11	2690.4	81.5	7259	18	US-10-263-153-30	Sequence 30, Appl
12	2690.4	81.5	7322	18	US-10-263-153-35	Sequence 35, Appl
13	2690.4	81.5	7352	18	US-10-263-153-25	Sequence 25, Appl
14	2690.4	81.5	7370	18	US-10-263-153-20	Sequence 20, Appl
15	2690.4	81.5	7370	18	US-10-263-153-61	Sequence 61, Appl
16	2690.4	81.5	7370	18	US-10-263-153-66	Sequence 66, Appl
17	2690.4	81.5	7370	18	US-10-263-153-71	Sequence 71, Appl
18	2690.4	81.5	7403	18	US-10-263-153-15	Sequence 15, Appl
19	2690.4	81.5	7442	18	US-10-263-153-10	Sequence 10, Appl
20	2690.4	81.5	7478	18	US-10-263-153-3	Sequence 3, Appli
21	2636.8	79.9	6806	18	US-09-908-943A-194	Sequence 194, App
22	2636.8	79.9	6806	20	US-10-801-487-194	Sequence 194, App
23	2636.8	79.9	6806	20	US-10-801-938-194	Sequence 194, App
24	2636.8	79.9	6806	20	US-10-801-509-194	Sequence 194, App
25	2636.8	79.9	6806	20	US-10-801-486-194	Sequence 194, App
26	2636.8	79.9	6806	21	US-10-801-493-194	Sequence 194, App
27	2605.4	79.0	7553	18	US-10-263-153-7	Sequence 7, Appli
28	1609.4	48.8	5558	16	US-10-241-596-137	Sequence 137, App
29	1474	44.7	4700	17	US-10-313-963A-54	Sequence 54, Appl
30	1462.4	44.3	4921	21	US-10-495-491-2	Sequence 2, Appli
31	1462.4	44.3	4935	21	US-10-495-491-1	Sequence 1, Appli
32	1462.4	44.3	4945	21	US-10-495-491-3	Sequence 3, Appli
33	1462.4	44.3	4951	21	US-10-495-491-4	Sequence 4, Appli
34	1417.6	43.0	1511	10	US-09-987-763-30	Sequence 30, Appl
35	1417.6	43.0	5926	13	US-10-024-809-3	Sequence 3, Appli
36	1238.4	37.5	5024	17	US-10-313-963A-52	Sequence 52, Appl
37	1238.4	37.5	5448	21	US-10-671-995A-1	Sequence 1, Appli
38	1238.4	37.5	5454	21	US-10-671-995A-2	Sequence 2, Appli
39	1237.2	37.5	1922	17	US-09-995-988A-32	Sequence 32, Appl
40	1237.2	37.5	1922	17	US-10-420-034A-32	Sequence 32, Appl
41	1237.2	37.5	6270	21	US-10-855-897-2	Sequence 2, Appli
42	1204	36.5	4557	17	US-10-457-372-5	GENERAL INFORMA
43	1204	36.5	6664	18	US-10-270-176-1	Sequence 1, Appli
44	1204	36.5	6668	18	US-10-270-176-2	Sequence 2, Appli
45	1204	36.5	6742	18	US-10-270-176-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-09-765-555-16
; Sequence 16, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m3 and zinc finger
; OTHER INFORMATION: protein 2Fpm3
US-09-765-555-16

Query Match 100.0%; Score 3300; DB 10; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCACACCATCGAATGGTCAAAACCTTTCCGGGTATGCGATGATAGCCCGGAAGAGA 60
Db 1 CGCACACCATCGAATGGTCAAAACCTTTCCGGGTATGCGATGATAGCCCGGAAGAGA 60
QY 61 GTCAAATTCAGGGTGGTGAATGTGAACACAGTAAACGGTTATACGATGTCGAGAGTATGCG 120
Db 61 GTCAAATTCAGGGTGGTGAATGTGAACACAGTAAACGGTTATACGATGTCGAGAGTATGCG 120
QY 121 GTGTCTCTTATCAGACCGGTTTCCGGGTGGTGAACCCAGGCGAGCCAGCTTTCTCGGAATA 180
Db 121 GTGTCTCTTATCAGACCGGTTTCCGGGTGGTGAACCCAGGCGAGCCAGCTTTCTCGGAATA 180
QY 181 CCGGGGAAAAGTGAAGCGGCGATGGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
Db 181 CCGGGGAAAAGTGAAGCGGCGATGGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGCGGGGCAAAACAGTGGTGTCTGATTTGGCGTGGCCACTTCCAGTCTGGGCCCTGC 300
Db 241 AACAACTGCGGGGCAAAACAGTGGTGTCTGATTTGGCGTGGCCACTTCCAGTCTGGGCCCTGC 300
QY 301 ACGCGCGTTCGCAAAATTTGTCGGGCGATTTAAATCTCGGCGCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGTTCGCAAAATTTGTCGGGCGATTTAAATCTCGGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTTCGATGTAGAACGAAGCGGCGTGAAGCGCTGTAAAGCGCGGTGCACAAATC 420
Db 361 TGGTGGTTCGATGTAGAACGAAGCGGCGTGAAGCGCTGTAAAGCGCGGTGCACAAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTCCGCGCGATCAACTGGGTGCCAGCG 480
Db 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTCCGCGCGATCAACTGGGTGCCAGCG 480
QY 481 TTGCTGTGGAAGTCCCTGCACATAATGTTCCGGGCTTATTTCTTGATGCTCTGACACAGA 540
Db 481 TTGCTGTGGAAGTCCCTGCACATAATGTTCCGGGCTTATTTCTTGATGCTCTGACACAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGGATCGGCGGTGGAGATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGGATCGGCGGTGGAGATC 600
QY 601 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGCTCGG 660
Db 601 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGCTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAG 720
QY 721 CGGAAACGGGAAGGCGACTGGAGTGCATGTCCGGTTTCAACAAACCATGCAATGCTGA 780
Db 721 CGGAAACGGGAAGGCGACTGGAGTGCATGTCCGGTTTCAACAAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCAACAGATCAGATGGCGTGGCGCAA 840
Db 781 ATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCAACAGATCAGATGGCGTGGCGCAA 840
QY 841 TGGCGGCATTTACCGAGTCCGGCTGCGGTTGCGGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGGCATTTACCGAGTCCGGCTGCGGTTGCGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTTATATCCCGCGTTTAAACCCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTTATATCCCGCGTTTAAACCCATCAACAGGATTTTC 960
QY 961 GCCTGTGGGGCAACCGAGGTGAACCGTTGTCGCAACTCTTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGTGGGGCAACCGAGGTGAACCGTTGTCGCAACTCTTCAGGGCCAGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGTGTAAGAAAACCAACCTTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGTGTAAGAAAACCAACCTTGGCGCCCAATA 1080
QY 1081 CGCAACCGGCTCTCCCGCGCGTGGCCGATTTAATTAATGCGAGTGGCACGACGATTT 1140

Db 1081 CGCAACCGGCTCTCCCGCGCGTGGCCGATTTAATTAATGCGAGTGGCACGACGATTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTGAACAGCTTATCATGCACTGCAAGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGAACAGCTTATCATGCACTGCAAGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACAGCCATCGGAAGCTGTGTATGGCTGTGACAGTGTGAATCACTGCAATAATTCG 1320
Db 1261 TCAGGACAGCCATCGGAAGCTGTGTATGGCTGTGACAGTGTGAATCACTGCAATAATTCG 1320
QY 1321 TGTGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCATAAACGGTT 1380
Db 1321 TGTGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCATAAACGGTT 1380
QY 1381 CTGGCAATATTTCTGAATGAGCTGTTGAACAATTAATCATCGGCTCGTATAATGTGTGGA 1440
Db 1381 CTGGCAATATTTCTGAATGAGCTGTTGAACAATTAATCATCGGCTCGTATAATGTGTGGA 1440
QY 1441 ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTGGTGTGTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTGGTGTGTTTTCACGA 1500
QY 1501 GCACTTCAACAAAGGACCATAGATTATGAAGCTGAAGAGGTAAACTGGTAAATCTCG 1560
Db 1501 GCACTTCAACAAAGGACCATAGATTATGAAGCTGAAGAGGTAAACTGGTAAATCTCG 1560
QY 1561 ATTAAACGGGCATTAAGGCTTAAACGCTCTCGCTGAAGTGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAAACGGGCATTAAGGCTTAAACGCTCTCGCTGAAGTGGTAAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGGTT 1680
QY 1681 GCGGCACTGGCGATGGCCCTGACATTTCTTCTGGGCAACGACCGCTTTGGTGCTAC 1740
Db 1681 GCGGCACTGGCGATGGCCCTGACATTTCTTCTGGGCAACGACCGCTTTGGTGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTGGCTGAAATCACCCGGAACAAAGCGTTCCAGACAGCTGAT 1800
Db 1741 GCTCAATCTGGCCTGTGGCTGAAATCACCCGGAACAAAGCGTTCCAGACAGCTGAT 1800
QY 1801 CGGTTTACTGGGATGCCGTACGTTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
Db 1801 CGGTTTACTGGGATGCCGTACGTTTACAACGGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTTACTGGGATTTTAAACAGATCTGCTGCCGAAACCCGCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTTACTGGGATTTTAAACAGATCTGCTGCCGAAACCCGCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGACGCGCTGATGTTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGACGCGCTGATGTTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGTAACGCGGTTATCGTTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGTAACGCGGTTATCGTTTCAAG 2040
QY 2041 TATGAAAACCGCAAGTACGACATTAAGAGCGTGGCGTGAATGCTGTAACGCGCGGAAGCG 2100
Db 2041 TATGAAAACCGCAAGTACGACATTAAGAGCGTGGCGTGAATGCTGTAACGCGCGGAAGCG 2100
QY 2101 GGTCTGACCTTCTCGGTGACCTGATTAATAAACAATGAATGCAACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGGTGACCTGATTAATAAACAATGAATGCAACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCGCGCTGG 2220

```
Db 2161 TCCATCGCAGAGCTGCCCTTTAATAAAGCGGAAACAGCGATGACCATCAACGGCCCGCTGG 2220
Qy 2221 GCATGGTCAACATCGACACACGAAAGTGAATTATGGTTAAACGGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACACGAAAGTGAATTATGGTTAAACGGTACTGCCGACCTTC 2280
Qy 2281 AAGGGTCAACCATCAAAACCGTTCTTGGCGTCTGAGCGCAGGTATTAAACGCGCCAGT 2340
Db 2281 AAGGGTCAACCATCAAAACCGTTCTTGGCGTCTGAGCGCAGGTATTAAACGCGCCAGT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACTATCTGCTCACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACTATCTGCTCACTGATGAAGTCTG 2400
Qy 2401 GAAGCGGTAAATAAGACAAACCGCTGGTCCGTAGCGCTGAAGTCTTACGAGAAAG 2460
Db 2401 GAAGCGGTAAATAAGACAAACCGCTGGTCCGTAGCGCTGAAGTCTTACGAGAAAG 2460
Qy 2461 TTGGCGAAGATCCACGTTATTGCCGCCACCATGGAAACGCCACGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGTTATTGCCGCCACCATGGAAACGCCACGAAAGGTGAATCATG 2520
Qy 2521 CCGAATCTCCGCAGATGTCGCTTCTTGGTATGCGTGGTACTGCGTGAATCAACGCC 2580
Db 2521 CCGAATCTCCGCAGATGTCGCTTCTTGGTATGCGTGGTACTGCGTGAATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACAGTCTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTCTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACACACACAAATAACAATACAAACCTCGGGATCGAGGGAAGGATTCAGAAATC 2700
Db 2641 AACACACACAAATAACAATACAAACCTCGGGATCGAGGGAAGGATTCAGAAATC 2700
Qy 2701 GGATCCTCTTCTCTGTGGCCAGCGGCCCTCGAGCCGGGAGAGCCCTATGCTGT 2760
Db 2701 GGATCCTCTTCTCTGTGGCCAGCGGCCCTCGAGCCGGGAGAGCCCTATGCTGT 2760
Qy 2761 CCGGAATGTGTAAGTCTTACAGCATCTGCGCCACCTGTTTCGACACGAGCGGCTCC 2880
Db 2761 CCGGAATGTGTAAGTCTTACAGCATCTGCGCCACCTGTTTCGACACGAGCGGCTCC 2880
Qy 2881 ACGGTTGAAAACCGTATAAATGSCCAGAGTGGCGGCAATCTTTAGCACACAGCGGCTCC 2880
Db 2881 ACGGTTGAAAACCGTATAAATGSCCAGAGTGGCGGCAATCTTTAGCACACAGCGGCTCC 2880
Qy 2881 CTGGTGGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCCAGATGTGGC 2940
Db 2881 CTGGTGGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCCAGATGTGGC 2940
Qy 2941 AAGTCTTTTCAGCCAGAGCTCCAGCCCTGGTGGCCACCAACGTAATCTCACACCGGGAGAG 3000
Db 2941 AAGTCTTTTCAGCCAGAGCTCCAGCCCTGGTGGCCACCAACGTAATCTCACACCGGGAGAG 3000
Qy 3001 CCCTATGCTTGTCCGGAATGTGGTAAGTCTTTACGCGCAGAGCAGTCCCTGTTGGCCAC 3060
Db 3001 CCCTATGCTTGTCCGGAATGTGGTAAGTCTTTACGCGCAGAGCAGTCCCTGTTGGCCAC 3060
Qy 3061 CAGCGTACCCACACGCGTGAAGACCGTATAATGCCAGAGTGGCGGAAATCTTTTAGT 3120
Db 3061 CAGCGTACCCACACGCGTGAAGACCGTATAATGCCAGAGTGGCGGAAATCTTTTAGT 3120
Qy 3121 GACTGCGCGACCTTGTGCGCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
Db 3121 GACTGCGCGACCTTGTGCGCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
Qy 3181 CAGAAATGTGGCAAGTCTTTTCTCCAAATCCAGCCATCTGCTCCGGACCAACGTAATCAT 3240
Db 3181 CAGAAATGTGGCAAGTCTTTTCTCCAAATCCAGCCATCTGCTCCGGACCAACGTAATCAT 3240
Qy 3241 ACCGTTAAAAAATAGTGGCCAGCGCGCAGTACCGGTACGAGTTCCGACCTACCGT 3300
Db 3241 ACCGTTAAAAAATAGTGGCCAGCGCGCAGTACCGGTACGAGTTCCGACCTACCGT 3300
```

RESULT 2

```
US-09-765-555-18
; Sequence 18, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Parial sequence of pMal-Ap3 and zinc finger
; OTHER INFORMATION: protein ZFPap3
US-09-765-555-18
```

Query Match 97.5%; Score 3216.8; DB 10; Length 3300;

Best Local Similarity 98.4%; Pred. No. 0;

Matches 3248; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

```
Qy 1 CCGACACCATCGAATGGTGCAGAAACCTTTTCGCGGTATGGCATGATAGCCCGCGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAGAAACCTTTTCGCGGTATGGCATGATAGCCCGCGAAGAGA 60
Qy 61 GTCAATTGAGGGTGGTGAATGTGAACCAAGTAACTTATACGATGTCGAGAGTATGCG 120
Db 61 GTCAATTGAGGGTGGTGAATGTGAACCAAGTAACTTATACGATGTCGAGAGTATGCG 120
Qy 121 GTGTCCTTATCAGACCTTTTCCCGCGTGTGAACACGAGCCAGCCACGCTTTCTCGGAAAA 180
Db 121 GTGTCCTTATCAGACCTTTTCCCGCGTGTGAACACGAGCCAGCCACGCTTTCTCGGAAAA 180
Qy 181 CCGGAAAAAAGTGAAGCGCGCATGGCGAGCTGAATTAATCCCAACCGCGTGGCAC 240
Db 181 CCGGAAAAAAGTGAAGCGCGCATGGCGAGCTGAATTAATCCCAACCGCGTGGCAC 240
Qy 241 AACAACTGCGGGGCAACAGTCTGTTGCTGATGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGCGGGGCAACAGTCTGTTGCTGATGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGGCGCTGCGCAATTCGCGCGGATTAATCTCGCGCGATCAATCTGGGTGCCAGCG 360
Db 301 ACGGCGCTGCGCAATTCGCGCGGATTAATCTCGCGCGATCAATCTGGGTGCCAGCG 360
Qy 361 TGGTGGTTCGATGTTAGAACGAGCGCGTGAAGCCCTGTAAGCGGCGGTGCAATTC 420
Db 361 TGGTGGTTCGATGTTAGAACGAGCGCGTGAAGCCCTGTAAGCGGCGGTGCAATTC 420
Qy 421 TTCTCGCGCAACGCGTCAAGTGGGTGATCAATTAATCTCGGTGATGATCAAGGATGCCA 480
Db 421 TTCTCGCGCAACGCGTCAAGTGGGTGATCAATTAATCTCGGTGATGATCAAGGATGCCA 480
Qy 481 TTGCTGTGGAAGTCCCTGCACTAATGTTCCGGCGTATTTCTTCTGATGTTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGTCCCTGCACTAATGTTCCGGCGTATTTCTTCTGATGTTCTGACCCAGA 540
Qy 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTCCGATGTTGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCGG 660
```


Db 601 |||||TGCTCGCATTTGGTACACGAAATCGCGCTGTGTAGCGGCCCAATTAAGTTCTGTCTCGG 660
Qy 661 CGCGTCTGGTCTGGCTGGCTGGCTGGAATAATATCTCTCTGCAATCAAAATTCAGCGGATAG 720
Db 661 CGCGTCTGGTCTGGCTGGCTGGCTGGAATAATATCTCTCTGCAATCAAAATTCAGCGGATAG 720
Qy 721 CGGAACGGGAAGCGCACTGGAGTGCATCTCGGTTTTTCAACAAACCAATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGCGCACTGGAGTGCATCTCGGTTTTTCAACAAACCAATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATCGTTCCCATCTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGCGCGCAA 840
Db 781 ATGAGGGCATCGTTCCCATCTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGCGCGCAA 840
Qy 841 TCGCGGCCAATACCGAGTCCGGGCTGGCGTTGGTGGCGGATATCTCGGTAGTGGGATAG 900
Db 841 TCGCGGCCAATACCGAGTCCGGGCTGGCGTTGGTGGCGGATATCTCGGTAGTGGGATAG 900
Qy 901 ACGATACCGAAGACAGCTCATGTATATCCCGCCGTTAAACACCAATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTATATCCCGCCGTTAAACACCAATCAAAACAGGATTTTC 960
Qy 961 GCCTGCTGGGCAACACGAGTGGACCGGTTGCTGCCAATCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGTGGACCGGTTGCTGCCAATCTCTCAGGGCCAGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGGCGGCTCTCACTGGTGAAGAAACACCACTGGCGGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCGGCTCTCACTGGTGAAGAAACACCACTGGCGGCCCAATA 1080
Qy 1081 CGCAACCGGCTCTCCCGCGGTTGGCCGATTCATTAATGCAGCTGSCACGACAGGTTT 1140
Db 1081 CGCAACCGGCTCTCCCGCGGTTGGCCGATTCATTAATGCAGCTGSCACGACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGAGTTAGTCACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGAGTTAGTCACTCATTAG 1200
Qy 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAAGGTGACCAATGCTTCTCGCG 1260
Db 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAAGGTGACCAATGCTTCTCGCG 1260
Qy 1261 TCAGGACGCAATCGAAGCTGTGTATGCTGTCGAGGTGTAATCACTGCAATAATTCG 1320
Db 1261 TCAGGACGCAATCGAAGCTGTGTATGCTGTCGAGGTGTAATCACTGCAATAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAATACCGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAATACCGTT 1380
Qy 1381 CTGGCAATATCTGAAATGAGCTGTTGCAAAATTAATCAATCATCGGCTGTAATAATGTTGGA 1440
Db 1381 CTGGCAATATCTGAAATGAGCTGTTGCAAAATTAATCAATCATCGGCTGTAATAATGTTGGA 1440
Qy 1441 ATTGTGAGCGGATAAACAATTTCAACAGGAAACAGCCAGTCCGTTTGGTGTGTTTACGA 1500
Db 1441 ATTGTGAGCGGATAAACAATTTCAACAGGAAACAGCCAGTCCGTTTGGTGTGTTTACGA 1500
Qy 1501 GCACTTCAACAAAGGACCATAGATATGAAACTGAAGAAGTAACTGGTAATCTCG 1560
Db 1501 GCACTTCAACAAAGGACCATAGATATGAAACTGAAGAAGTAACTGGTAATCTCG 1560
Qy 1561 ATTAACGGGATTAAGGCTATACGTTCTCGTGAAGTCGGTAAGAAATTCGAGAAGAT 1620
Db 1561 ATTAACGGGATTAAGGCTATACGTTCTCGTGAAGTCGGTAAGAAATTCGAGAAGAT 1620
Qy 1621 ACCGGAATTAAGTACCGTTGAGCATCCGGATAAATCTGGAAGAAATTCGCCAGGTT 1680
Db 1621 ACCGGAATTAAGTACCGTTGAGCATCCGGATAAATCTGGAAGAAATTCGCCAGGTT 1680
Qy 1681 GCGGCAATCGGAGTGGCCCTGACATTAATCTTGGGCAACAGCCGCTTTGGTGGCTAC 1740

Db 1681 GCGGCAATCGGAGTGGCCCTGACATTAATCTTCTGGGCACACGACCGGTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGACAAAGCGTTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGACAAAGCGTTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTAACAACGCAAGCTGATGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTAACAACGCAAGCTGATGCTTACCCGATCGCTGTT 1860
Qy 1861 GAACGCTTATCGCTGATTTATTAACAAGATCTGTCGCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAACGCTTATCGCTGATTTATTAACAAGATCTGTCGCGAACCCGCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGGTACTTCACTGCGCGCTGATGCTGCTGACGCGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGGTACTTCACTGCGCGCTGATGCTGCTGACGCGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTAAGAGCGTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGAGCGTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGTTGTAACCTGATTAATAAACAACAATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTGTAACCTGATTAATAAACAACAACAATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAAGCTGCTTTTAATGAAGCGAAACAGCGATGACCATCAACGCCCGTGG 2220
Db 2161 TCCATCGCAAGCTGCTTTTAATGAAGCGAAACAGCGATGACCATCAACGCCCGTGG 2220
Qy 2221 GCATGCTCAACATGCACACACGAAAGTGAATTAATGCTGTAACGGTAATGCGCGACCTTC 2280
Db 2221 GCATGCTCAACATGCACACACGAAAGTGAATTAATGCTGTAACGGTAATGCGCGACCTTC 2280
Qy 2281 AAGGCTCAACCATCCAAACCGTTGCTGGCTGCTGAGCGCAGGTATTAACCGCGCAGT 2340
Db 2281 AAGGCTCAACCATCCAAACCGTTGCTGGCTGCTGAGCGCAGGTATTAACCGCGCAGT 2340
Qy 2341 CCGAAACAAAGAGCTGCAAAAGAGTTCTCGAATACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGCAAAAGAGTTCTCGAATACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAACGCTTAAATAAGACAAAACCGCTGGTGGCTAGCGCTGAAGTCTTAACGAGGAAGAG 2460
Db 2401 GAACGCTTAAATAAGACAAAACCGCTGGTGGCTAGCGCTGAAGTCTTAACGAGGAAGAG 2460
Qy 2461 TTGGCGAAAGATCCAGTATTCGCCACCATGGAAGACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTCGCCACCATGGAAGACGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACAATCCCGCAGATGTCGCTTTCTGSTATGCGTGGCTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACAATCCCGCAGATGTCGCTTTCTGSTATGCGTGGCTGCTGCTGATCAACGCC 2580
Qy 2581 GCGAGGCTGCTGAGACTGTCGATGAAGCTGTAAGAGCGCGCAGACTAAATTCGAGTCTG 2640
Db 2581 GCGAGGCTGCTGAGACTGTCGATGAAGCTGTAAGAGCGCGCAGACTAAATTCGAGTCTG 2640
Qy 2641 AAACAACAACAATAAACAATAACAACCTCGGATCGAGGGAAGGATTTCAAGATTC 2700
Db 2641 AAACAACAACAATAAACAATAACAACCTCGGATCGAGGGAAGGATTTCAAGATTC 2700
Qy 2701 GATCCTCTTCTGTCGCGCAGCGCCCTCGAGCCCGGGGAAGCCCTATCTCTTGT 2760
Db 2701 GATCCTCTTCTGTCGCGCAGCGCCCTCGAGCCCGGGGAAGCCCTATCTCTTGT 2760
Qy 2761 CCGGAATGTGGTAAGTCTTCAAGCATCTCGGCAACCTGGTTTCGCCACAGGCTACCCAC 2820
Db 2761 CCGGAATGTGGTAAGTCTTCAAGCATCTTCCCTGGTGGTGGCCACAGGCTACCCAC 2820

Db 1201 GCACAAATTCATGTTTGACAGCTTATCATCGACTGCAAGGTGCAACCAATGCTTCTGGCG 1260
QY 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGAGGTGCTGTAATCACTGCATAAATTCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGAGGTGCTGTAATCACTGCATAAATTCG 1320
QY 1321 TGTGCTCAAGGCGCACTCCGTTCTGGATAATGTTTTTGGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGGCGCACTCCGTTCTGGATAATGTTTTTGGCGCGACATCAATACGGTT 1380
QY 1381 CTGGCAATATTTCTGAATAGCTGTGACATTAATCATCGGCTCGTATTAATGTTGGA 1440
Db 1381 CTGGCAATATTTCTGAATAGCTGTGACATTAATCATCGGCTCGTATTAATGTTGGA 1440
QY 1441 ATTGTAGCGGATAACAATTTTACACAGGAAACAGCCAGTCCGTTTAGTGTGTTTCAAGA 1500
Db 1441 ATTGTAGCGGATAACAATTTTACACAGGAAACAGCCAGTCCGTTTAGTGTGTTTCAAGA 1500
QY 1501 GCACCTCACCAACAGGACCATAGATTATGAAAACCTGAAGAGGTAAACTGGTAAATCTGG 1560
Db 1501 GCACCTCACCAACAGGACCATAGATTATGAAAACCTGAAGAGGTAAACTGGTAAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAAACGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAAACGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCGCCAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCGCCAGGTT 1680
QY 1681 GCGCAACTGGCGATGCGCTGACATTAATCTTCTGGGACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGCGCTGACATTAATCTTCTGGGACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTTGGCTGAAATCAACCCGGACAAAGCGTTCAGGACAAAGCTGAT 1800
Db 1741 GCTCAATCTGGCTGTTGGCTGAAATCAACCCGGACAAAGCGTTCAGGACAAAGCTGAT 1800
QY 1801 CCGTTTACCTGGATGCGGTACGTTTAAACGCGAAGCTGATGCTTACCGGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGGTACGTTTAAACGCGAAGCTGATGCTTACCGGATCGCTGTT 1860
QY 1861 GAAGGTTTATCGCTGATTATTAACAAGATCTGCTGCGGAAACGCGGCTGATGTTCAAC 1920
Db 1861 GAAGGTTTATCGCTGATTATTAACAAGATCTGCTGCGGAAACGCGGCTGATGTTCAAC 1920
QY 1921 GAGATCCGCGCTGGATAAAGAACTGAAACGGAAGGTAAAGCGGCTGATGTTCAAC 1980
Db 1921 GAGATCCGCGCTGGATAAAGAACTGAAACGGAAGGTAAAGCGGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACGTTACTTCACTGCGGCTGATGCTGCTGACGCGGCTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACGTTACTTCACTGCGGCTGATGCTGCTGACGCGGCTTATGCGTTCAAG 2040
QY 2041 TATGAAAACGCGCAAGTAGTACGATTAAGACGTTGGGCTGGATACGCTGGCGGAAGCG 2100
Db 2041 TATGAAAACGCGCAAGTAGTACGATTAAGACGTTGGGCTGGATACGCTGGCGGAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAAACAAACACATGATGATGATGATGATGAT 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAAACAAACACATGATGATGATGATGATGAT 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAAGAGGGAACACGCGATGACCATCAACGCGCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAAGAGGGAACACGCGATGACCATCAACGCGCGCTGG 2220
QY 2221 GCATGGTCCAACTCGACACAGCAAAAGTAAATTAATGCTGTAACGCTGCTGCGACCTTC 2280
Db 2221 GCATGGTCCAACTCGACACAGCAAAAGTAAATTAATGCTGTAACGCTGCTGCGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTGTTGGGCTGCTGAGCGCAGGTATTAACGCGCGCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTGTTGGGCTGCTGAGCGCAGGTATTAACGCGCGCAGT 2340

RESULT 4

US-09-765-555-15

; Sequence 15, Application US/09765555

; Publication No. US20030037355A1

; GENERAL INFORMATION:

; APPLICANT: The Scripps Research Institute

; TITLE OF INVENTION: Methods and compositions to modulate

;
; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m2 and zinc finger
; OTHER INFORMATION: protein ZFPm2
US-09-765-555-15

Query Match 97.2%; Score 3207.2; DB 10; Length 3300;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY	1	CGGACACCATCAAAATGGTGC	AAACCTTTTCGGGGTATGGCATGATAGCCCGGAAGAGA	60
DB	1	CGGACACCATCAAAATGGTGC	AAACCTTTTCGGGGTATGGCATGATAGCCCGGAAGAGA	60
QY	61	GTCAATTACAGGTGTGAATGT	GAACACAGTAACGTTATACGATGTCGCAGAGTATGCCG	120
DB	61	GTCAATTACAGGTGTGAATGT	GAACACAGTAACGTTATACGATGTCGCAGAGTATGCCG	120
QY	121	GTGTCTCTTATCAGACCGTT	CCCCGGTGTGAACAGCCAGCCACGCTTTCTCGAAAA	180
DB	121	GTGTCTCTTATCAGACCGTT	CCCCGGTGTGAACAGCCAGCCACGCTTTCTCGAAAA	180
QY	181	CGCGGAAAAAGTGAAGCGG	CGATGGCGAGCTGAATTACATTCCTCAACCCGCTGGCAC	240
DB	181	CGCGGAAAAAGTGAAGCGG	CGATGGCGAGCTGAATTACATTCCTCAACCCGCTGGCAC	240
QY	241	AACAACTGGCGGCAAAAC	AGTCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
DB	241	AACAACTGGCGGCAAAAC	AGTCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
QY	301	ACGCGCGTTCGAAATGT	CGCGCGAATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360
DB	301	ACGCGCGTTCGAAATGT	CGCGCGAATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360
QY	361	TGGTGTGTCGATGTAGAAC	GAAGCGGCGTGAAGCGGCTGTAAAGCGGCGTGCACAATC	420
DB	361	TGGTGTGTCGATGTAGAAC	GAAGCGGCGTGAAGCGGCTGTAAAGCGGCGTGCACAATC	420
QY	421	TTCTCGGCAACCGTTCAG	TGGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCCA	480
DB	421	TTCTCGGCAACCGTTCAG	TGGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCCA	480
QY	481	TTGCTGTGGAAGCTGCT	GCATTAATGTTCCGGCGTTATTTCTTGATGTCCTGACACAGA	540
DB	481	TTGCTGTGGAAGCTGCT	GCATTAATGTTCCGGCGTTATTTCTTGATGTCCTGACACAGA	540
QY	541	CACCCATCAACAGTATTA	TTTCTCCATGAAGACGTTACGCGACTGGGCGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTA	TTTCTCCATGAAGACGTTACGCGACTGGGCGTGGAGCATC	600
QY	601	TGGTGCATTTGGTGCAC	AGAAATCGCTGTAGCGGGCCCATTAAGTCTGTCTCGG	660
DB	601	TGGTGCATTTGGTGCAC	AGAAATCGCTGTAGCGGGCCCATTAAGTCTGTCTCGG	660
QY	661	CGCGTCTGGCTGGCTGC	ATATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	661	CGCGTCTGGCTGGCTGC	ATATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAACGGGAAGCGCAT	GTGGATGCGATTCGGTTTTCACAAACCATGCAAAATGCTGA	780
DB				

DB	721	CGGAACGGGAAGCGCAT	GTGGATGCGATGTCTGGTTTTCACAAACCATGCAAAATGCTGA	780
QY	781	ATGAGGCGCATCGTTCC	CACTGCGATGCTGGTTGGCAACGATCAGATGCGCTGGCGCGAA	840
DB	781	ATGAGGCGCATCGTTCC	CACTGCGATGCTGGTTGGCAACGATCAGATGCGCTGGCGCGAA	840
QY	841	TGCGGCGCATTAACGAG	TCCGGCTCGCGGATATCTCGGTAGTGGGATAGC	900
DB	841	TGCGGCGCATTAACGAG	TCCGGCTCGCGGATATCTCGGTAGTGGGATAGC	900
QY	901	ACGATACCGAAGACAG	CTCATGTTATATCCCGCGTTAAACCAACCATCAAAACGAGATTTC	960
DB	901	ACGATACCGAAGACAG	CTCATGTTATATCCCGCGTTAAACCAACCATCAAAACGAGATTTC	960
QY	961	GCCTGCTGGGGCAAA	ACACGAGCTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGTGA	1020
DB	961	GCCTGCTGGGGCAAA	ACACGAGCTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGTGA	1020
QY	1021	AGGGCAATCAGCTGTT	GCCCGTCTCACTGGTGAAGAAAAACCACTGGCGGCCCAATA	1080
DB	1021	AGGGCAATCAGCTGTT	GCCCGTCTCACTGGTGAAGAAAAACCACTGGCGGCCCAATA	1080
QY	1081	CGCAAAACCGCTCTCC	CGCGGTTGGCCGATTCAATTAATGCAGCTGGCAGACAGGTTT	1140
DB	1081	CGCAAAACCGCTCTCC	CGCGGTTGGCCGATTCAATTAATGCAGCTGGCAGACAGGTTT	1140
QY	1141	CCGACTGGAAGGGG	CAGTCAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200
DB	1141	CCGACTGGAAGGGG	CAGTCAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200
QY	1201	GCACAAATCTCATGTT	TGACAGCTTATCATCGCTGCAACCAATGCTTCTGCGG	1260
DB	1201	GCACAAATCTCATGTT	TGACAGCTTATCATCGCTGCAACCAATGCTTCTGCGG	1260
QY	1261	TCAGCAGCCATCGGA	AGCTGTGATGCTGTGAGCTGTAATCACTGCATTAATTCG	1320
DB	1261	TCAGCAGCCATCGGA	AGCTGTGATGCTGTGAGCTGTAATCACTGCATTAATTCG	1320
QY	1321	TGTGCTCAAGCGCA	CTCCCGTTCTGGATTAATGTTTTGCGCGCAATCAATACGGTT	1380
DB	1321	TGTGCTCAAGCGCA	CTCCCGTTCTGGATTAATGTTTTGCGCGCAATCAATACGGTT	1380
QY	1381	CTGGCAAAATCTGAA	ATAGCTGTTGACATTAATCATCGCTCGTATTAATGTGTGA	1440
DB	1381	CTGGCAAAATCTGAA	ATAGCTGTTGACATTAATCATCGCTCGTATTAATGTGTGA	1440
QY	1441	ATTGTGAGCGGAT	TAACAAATTTTACACAGGAACAGCAGTCCGTTTAGTGTTCACGA	1500
DB	1441	ATTGTGAGCGGAT	TAACAAATTTTACACAGGAACAGCAGTCCGTTTAGTGTTCACGA	1500
QY	1501	GCACCTTCAACCA	AGGACCATAGATTATGAAACTGGAAGAGTAAACTTGGTAATCTGG	1560
DB	1501	GCACCTTCAACCA	AGGACCATAGATTATGAAACTGGAAGAGTAAACTTGGTAATCTGG	1560
QY	1561	ATTAAACCGCGATA	AAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT	1620
DB	1561	ATTAAACCGCGATA	AAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAAGT	CACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCACAGGTT	1680
DB	1621	ACCGGAATTAAGT	CACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCACAGGTT	1680
QY	1681	GGCGCAACTGCG	ATGAGCTTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
DB	1681	GGCGCAACTGCG	ATGAGCTTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGCG	CTGTTGGCTGAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGAT	1800
DB	1741	GCTCAATCTGCG	CTGTTGGCTGAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGAT	1800
QY	1801	CGGTTTACCTGG	ATGCGTGTACAAACCGCAAGCTGATGCTTACCCGATCCCTGTT	1860
DB	1801	CGGTTTACCTGG	ATGCGTGTACAAACCGCAAGCTGATGCTTACCCGATCCCTGTT	1860

```
QY 1861 GAAGCGTTATCGTGATTTTAAACAAAGATCTGTCGCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGTGATTTTAAACAAAGATCTGTCGCGAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGTGAAGAGCGCGCTGATTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGTGAAGAGCGCGCTGATTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTACCTCGCGCTGATTTAAAGACGTGGCGTGGATTAACGCTGGCGCGAAAGCG 2040
Db 1981 CTGCAAGAACCGTACTTACCTCGCGCTGATTTAAAGACGTGGCGTGGATTAACGCTGGCGCGAAAGCG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGCTGATTAAGAACTGAAGCGTGGCGTGGATTAACGCTGGCGCGAAAGCG 2160
Db 2101 GGTCTGACCTTCTGCTGATTAAGAACTGAAGCGTGGCGTGGATTAACGCTGGCGCGAAAGCG 2160
QY 2161 TCCATCGCAGAGCTGCGCTTTTAAAGCGGAAACAGCGATGACCATCAACGGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCGCTTTTAAAGCGGAAACAGCGATGACCATCAACGGCCGCTGG 2220
QY 2221 GCATGGTCAACATCGACACCGAAGAGTGAATTAAGTGAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACCGAAGAGTGAATTAAGTGAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGTCACACCATCAACCGTTCGTCGCGTGTGAGCGAGGTATTAACGGCGCAGT 2340
Db 2281 AAGGTCACACCATCAACCGTTCGTCGCGTGTGAGCGAGGTATTAACGGCGCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
QY 2461 TTGGCGAAAGATCCACGCTATTGCGCCACCATGGAACCGCCAGAACGCTGAAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGCTATTGCGCCACCATGGAACCGCCAGAACGCTGAAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
QY 2581 GCCAGCGTCTGACACTGTCGATGACGCTTCTGGTATGCGCTGCTGCTGCTGCTGCTGCTGCTG 2640
Db 2581 GCCAGCGTCTGACACTGTCGATGACGCTTCTGGTATGCGCTGCTGCTGCTGCTGCTGCTGCTG 2640
QY 2641 AACAAACAAACAAATAACAAACAAACCTCGGGATCGAGGGAAGGATTTTCAAGATTC 2700
Db 2641 AACAAACAAACAAATAACAAACAAACCTCGGGATCGAGGGAAGGATTTTCAAGATTC 2700
QY 2701 GGATCTCTTCTCTGTGGCCAGCGGCTCTGAGCGCGGGGAGAGCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTGGCCAGCGGCTCTGAGCGCGGGGAGAGCCTATGCTTGT 2760
QY 2761 CCGGAATGTGTAGTCTCTCAGCGATCTGGCCACCTGGTTCGCCACGAGCTGATCCAC 2820
Db 2761 CCGGAATGTGTAGTCTCTCAGCGATCTGGCCACCTGGTTCGCCACGAGCTGATCCAC 2820
QY 2821 ACGGGTGAACAAACCGTATAAATGCCAGAGTGGCGCAAACTTTTAGCACACGCGCTCC 2880
Db 2821 ACGGGTGAACAAACCGTATAAATGCCAGAGTGGCGCAAACTTTTAGCACACGCGCTCC 2880
QY 2881 CTGGTGGCCATCAACGCTCTCATCTGCGGAGGAGCCATACAAATGTCAGAAATGTGGC 2940
Db 2881 CTGGTGGCCATCAACGCTCTCATCTGCGGAGGAGCCATACAAATGTCAGAAATGTGGC 2940
```

```
QY 2941 AAGTCTTTACCCAGAGCTCAGCGCTGGTGGCCCAACAAAGTACTCTACACCGGGGAGAAG 3000
Db 2941 AAGTCTTTCTCGGTCTGCAATCTCGTCCGGCAACAAAGTACTCTACACCGGGGAGAAG 3000
QY 3001 CCCTATGCTTCTCGGGAATGTGAAGTCTCTTACGCCAGAGCAGCTCCCTGGTGGCCAC 3060
Db 3001 CCCTATGCTTCTCGGGAATGTGAAGTCTCTTACGCCAGAGCAGCTCCCTGGTGGCCAC 3060
QY 3061 CAGCGTACCCACACCGGCGGAAACCCGTATAAATGAGTGGCGGAGCCATACAAATGT 3120
Db 3061 CAGCGTACCCACACCGGCGGAAACCCGTATAAATGAGTGGCGGAGCCATACAAATGT 3120
QY 3121 GACTGCCGCGACCTTCTCGGCTCAACGCTCACTATGCTGGCGGAGCCATACAAATGT 3180
Db 3121 GACTGCCGCGACCTTCTCGGCTCAACGCTCACTATGCTGGCGGAGCCATACAAATGT 3180
QY 3181 CCAGAATGTGGCAAGTCTTTCTCCCAATCCAGCCATCTCGTCCGGCAACATCTCGTCCGGCAACGCTACTCAC 3240
Db 3181 CCAGAATGTGGCAAGTCTTTCTCGGCTGCAACATCTCGTCCGGCAACGCTACTCAC 3240
QY 3241 ACCGGTAAAAAATACTAGTGGCGCGCGCGCGAGTACCGGTACGAGTTCGGGACTACGCT 3300
Db 3241 ACCGGTAAAAAATACTAGTGGCGCGCGCGCGAGTACCGGTACGAGTTCGGGACTACGCT 3300
```

```
RESULT 5
US-09-765-555-17
; Sequence 17, Application US/09765555
; Publication No. US2003003735A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m4 and zinc finger
; OTHER INFORMATION: protein ZFPm4
US-09-765-555-17
```

```
Query Match 97.2%; Score 3207.2; DB 10; Length 3300;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTTCGGGGTATGCGATGATGCGCGGAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAGAAACCTTTTCGGGGTATGCGATGATGCGCGGAGAGA 60
QY 61 GTCAATTTCAGGTGGTGAATGTGAACCAAGTAACTGATGATGATGCGAGATGCGG 120
Db 61 GTCAATTTCAGGTGGTGAATGTGAACCAAGTAACTGATGATGATGCGAGATGCGG 120
QY 121 GTGCTCTTTATCAGACCGGTTTCCCGGTGTGAATGTGAACCAAGTAACTGATGATGCGAGATGCGG 180
Db 121 GTGCTCTTTATCAGACCGGTTTCCCGGTGTGAATGTGAACCAAGTAACTGATGATGCGAGATGCGG 180
QY 181 CCGGGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTTACATTTCCCAACCGCGTGGC 240
Db 181 CCGGGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTTACATTTCCCAACCGCGTGGC 240
QY 241 AACAACTGGCGGCAAAACAGTCTGCTGATTTGGGCTGGCACTCCAGTCTGGCCCTGC 300
```

Db 241 AACAACTGGCGGCAAA CAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCGCGCCCTGC 300
Qy 301 AGCGCGCTCGCAAAATGTCGCGCGAATAAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
Db 301 AGCGCGCTCGCAAAATGTCGCGCGAATAAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
Qy 361 TGGTGGTGCAGTGTAGAAAGCGGCGTCGAGCGCTGTAAAGCGGCGTGCACAATC 420
Db 361 TGGTGGTGCAGTGTAGAAAGCGGCGTCGAGCGCTGTAAAGCGGCGTGCACAATC 420
Qy 421 TTCTCGCGCAACGCGTCAGTGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCCA 480
Db 421 TTCTCGCGCAACGCGTCAGTGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCCCTGCAATAATGTTCCGCGCTTATTTCTGATGCTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCCCTGCAATAATGTTCCGCGCTTATTTCTGATGCTCTGACCCAGA 540
Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGCTACCGGCTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGCTACCGGCTGGGCGTGGAGCATC 600
Qy 601 TGGTGGCATTTGGTCAACAGCAAAATCGGCTGTTAGCGGCGCCATTAAGTTCGTCTCGG 660
Db 601 TGGTGGCATTTGGTCAACAGCAAAATCGGCTGTTAGCGGCGCCATTAAGTTCGTCTCGG 660
Qy 661 CGCGTCTCGCTGCGTGGCTGCGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTGCGTGGCTGCGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTCGCAACCAACCAATTCGCAATGCTGA 780
Db 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTCGCAACCAACCAATTCGCAATGCTGA 780
Qy 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTCGCAACGATCAGATGGCGTGGCGGCA 840
Db 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTCGCAACGATCAGATGGCGTGGCGGCA 840
Qy 841 TCGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATAG 900
Db 841 TCGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATAG 900
Qy 901 AGGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAAAAGGATTTTC 960
Db 901 AGGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAAAAGGATTTTC 960
Qy 961 GCCTGCTGGGCAACACCGAGTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACCGAGTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAAACCACTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAAACCACTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCTCTCCCGCGGTTGGCGATTCATTAATGAGCTGGCACACAGGTTT 1140
Db 1081 CGCAACCGCTCTCCCGCGGTTGGCGATTCATTAATGAGCTGGCACACAGGTTT 1140
Qy 1141 CCCGACTGGGAAGCGGCGAGTGCAGCGCAACCAATTAATGTCAGTTAGCTCACTCATTAG 1200
Db 1141 CCCGACTGGGAAGCGGCGAGTGCAGCGCAACCAATTAATGTCAGTTAGCTCACTCATTAG 1200
Qy 1201 GCACAATCTCATGTTTGACAGCTTATCATGACTGCAGGTCGACCAATGCTTCTGGCG 1260
Db 1201 GCACAATCTCATGTTTGACAGCTTATCATGACTGCAGGTCGACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCGTGTGAGGTCGTAATCACTGCATTAATTCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCGTGTGAGGTCGTAATCACTGCATTAATTCG 1320
Qy 1321 TGTGCTCAGGCGGCACTCCCGTCTGGATATGTTTTTGGCGGACATCATACGGTT 1380
Db 1321 TGTGCTCAGGCGGCACTCCCGTCTGGATATGTTTTTGGCGGACATCATACGGTT 1380

Qy 1381 CTGGCAAAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATATGTTGGA 1440
Db 1381 CTGGCAAAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATATGTTGGA 1440
Qy 1441 ATTGTGACGGATACCAATTTTCACACAGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
Db 1441 ATTGTGACGGATACCAATTTTCACACAGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
Qy 1501 GCACCTTCAACCAAGGACCAATAGATTATGAAAACTGAAGAGGTAAACTGGTAATCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCAATAGATTATGAAAACTGAAGAGGTAAACTGGTAATCTGG 1560
Qy 1561 ATTAACCGCGATAAAGGCTATAACCGTCTCGTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACCGCGATAAAGGCTATAACCGTCTCGTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTACACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCCTCCAGGTT 1680
Db 1621 ACCGGAATTAAGTACACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCCTCCAGGTT 1680
Qy 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTAGCTTACAAACCGCAAGCTGTTTACCCGATCCCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTAGCTTACAAACCGCAAGCTGTTTACCCGATCCCTGTT 1860
Qy 1861 GAAGGCTTATCGCTGATTATAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAAGGCTTATCGCTGATTATAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCGCTGGATAAGAACTGAAAGCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAGAACTGAAAGCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGCTGAGCGGGGTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGCTGAGCGGGGTTATCGGTTCAAG 2040
Qy 2041 TATGAAAAACGCAAGTACGACATTAAGACAGTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGACAGTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGGTTGACCTGATTAAAAACCAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGGTTGACCTGATTAAAAACCAACACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGAGAAAGCTGCGCTTTAAAGCGGAAACAGGATGACCATCAACGGCCGCTGG 2220
Db 2161 TCCATCGAGAAAGCTGCGCTTTAAAGCGGAAACAGGATGACCATCAACGGCCGCTGG 2220
Qy 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAATGTTGTAACCGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAATGTTGTAACCGTACTGCGGACCTTC 2280
Qy 2281 AAGGTCACCAATCCAAACCGTTCGTTGGCGTGTGAGCGGAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTCACCAATCCAAACCGTTCGTTGGCGTGTGAGCGGAGGTATTAACGCCGCCAGT 2340
Qy 2341 CGGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
Db 2341 CGGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCCGTAGCGCTGAACTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCCGTAGCGCTGAACTCTTACGAGGAAGAG 2460

QY 541 CACCCATCAACAGTATTATTTCTCCATGAAGCGGTACGGCACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCATGAAGCGGTACGGCACTGGCGGTGGAGCATC 600
QY 601 TGGTCGCAATTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAAGTTCTGTCTCGG 660
Db 601 TGGTCGCAATTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGCTGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCACTGGAAGTCGCATGTCCGGTTTTTCAACAACCAATGCAATGCTGA 780
Db 721 CGGAACGGGAAGGCACTGGAAGTCGCATGTCCGGTTTTTCAACAACCAATGCAATGCTGA 780
QY 781 ATGAGGGCATCGTTCGCCACTGGATGCTGGTTGCCAACAGATGGCGCTGGGCGCAA 840
Db 781 ATGAGGGCATCGTTCGCCACTGGATGCTGGTTGCCAACAGATGGCGCTGGGCGCAA 840
QY 841 TGGCGGCCATTACCGAGTCCGGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATAG 900
Db 841 TGGCGGCCATTACCGAGTCCGGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTATATCCCGCCGTTAACCHACCATCAAA CHGAGTTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTATATCCCGCCGTTAACCHACCATCAAA CHGAGTTTTC 960
QY 961 GCCTCTGGGGCAACACGAGCTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTCTGGGGCAACACGAGCTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAGAAACCAACCCCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAGAAACCAACCCCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGAATTAATGAGCTGGCAACAGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGAATTAATGAGCTGGCAACAGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGAATTAATGTAGTTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGAATTAATGTAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAATTCATGTTTGACAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTCATGTTTGACAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTCGTAATCACTGCATATTCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTCGTAATCACTGCATATTCG 1320
QY 1321 TGTCTCTCAAGCGCACTCCCGTCTGGATATGTTTTTGGCGGACATCAATACGTT 1380
Db 1321 TGTCTCTCAAGCGCACTCCCGTCTGGATATGTTTTTGGCGGACATCAATACGTT 1380
QY 1381 CTGGCAAAATATCTGAAATGAGCTGTGACAATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAAAATATCTGAAATGAGCTGTGACAATTAATCATCGGCTCGTATATGTTGGA 1440
QY 1441 ATTGTGAGCGGATAACAAATTCACAGAAACAGCCAGTCCGTTAGGTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTCACAGAAACAGCCAGTCCGTTAGGTTTTCACGA 1500
QY 1501 GCACCTCACCAACAGGACCATAGATTATGAAAACCTGAAGAGGTAAATCTGATCTGG 1560
Db 1501 GCACCTCACCAACAGGACCATAGATTATGAAAACCTGAAGAGGTAAATCTGATCTGG 1560
QY 1561 ATTTAAGCGGATAAAGGCTTAAACCGTCTCGCTGAAAGTCCGTTAGGTTTTCAGAAAGAT 1620
Db 1561 ATTTAAGCGGATAAAGGCTTAAACCGTCTCGCTGAAAGTCCGTTAGGTTTTCAGAAAGAT 1620
QY 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAACTGGAAAGAGAAATTCACAGGTT 1680

Db 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGATAACTGGAAGAGAAATTCACAGGTT 1680
QY 1681 CGGGCAACTGGCGATGGCCCTGACATATCTTTCTGGGCACACGACCGCTTTGGTGGCTTAC 1740
Db 1681 CGGGCAACTGGCGATGGCCCTGACATATCTTTCTGGGCACACGACCGCTTTGGTGGCTTAC 1740
QY 1741 GCTCAATCTGGCCTGTGGCTGGAATCAACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTGGCTGGAATCAACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
QY 1801 CGGTTTACCTGGGATGGCTAGCTTACAAACGCAAGCTGATTTGCTTACCCGATCCGCTGT 1860
Db 1801 CGGTTTACCTGGGATGGCTAGCTTACAAACGCAAGCTGATTTGCTTACCCGATCCGCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATATAAGAACTGAAAGGAAAGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATATAAGAACTGAAAGGAAAGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGGCAATGACATTAAGAACGCTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGGCAATGACATTAAGAACGCTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACATGAAATGCGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACATGAAATGCGAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCGCTTTAATAAGCGGCAACAGCGATGACCATCAACGGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCGCTTTAATAAGCGGCAACAGCGATGACCATCAACGGCCGCTGG 2220
QY 2221 GCATGCTCAACATCGACCAACGAAAGTGAATTAATGCTGTAACCGTACTGCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACCAACGAAAGTGAATTAATGCTGTAACCGTACTGCGGACCTTC 2280
QY 2281 AAGGTCACACCATCAAAACCGTTGTTGGCGTGTGAGCGAGGTTAATAAGCGCCCACT 2340
Db 2281 AAGGTCACACCATCAAAACCGTTGTTGGCGTGTGAGCGAGGTTAATAAGCGCCCACT 2340
QY 2341 CGGAACAAAGAGCTGGCAAAAGAGTTCTCGAATAACTATCTGCTGACTGATGAAGGCTCG 2400
Db 2341 CGGAACAAAGAGCTGGCAAAAGAGTTCTCGAATAACTATCTGCTGACTGATGAAGGCTCG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCACGATTTGCGGCCACCATATGGAACACGCGCCAGAAAGGTTGAAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTTGCGGCCACCATATGGAACACGCGCCAGAAAGGTTGAAATCATG 2520
QY 2521 CGGAACATCCGCGAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCG 2580
Db 2521 CGGAACATCCGCGAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCG 2580
QY 2581 GCCAGCGGTGCTGACATGTCGATGAAGCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGCTGACATGTCGATGAAGCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACCAACAAACAAATAACAAACAACTCCGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACCAACAAACAAATAACAAACAACTCCGGATCGAGGGAAGGATTTTCAGAAATTC 2700
QY 2701 GGATCCTCTTCTGCTGGCCGAGCGGCTCTGAGCGCGG 2741

Db 2701 GGATCCTCAGTCCTCCGGGCAAGCCGCGCGCTGGTGGGAGG 2741

RESULT 7
US-10-149-472-5
; Sequence 5 Application US/10149472
; Publication No. US20040029204A1
; GENERAL INFORMATION:
; APPLICANT: DOUBIN-GRAMATICA, FRANCOISE
; APPLICANT: DOUBIN-GRAMATICA, FRANCOISE
; APPLICANT: GREGOIRE
; APPLICANT: GREGOIRE
; TITLE OF INVENTION: METHOD FOR OBTAINING HUMAN CDC25 PHOSPHATASES AND METHOD
; TITLE OF INVENTION: FOR IDENTIFYING HUMAN CDC25 PHOSPHATASE MODULATORS
; FILE REFERENCE: bml-427.065
; CURRENT APPLICATION NUMBER: US/10/149,472
; CURRENT FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: PCT/FR00/03496
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: FR 99/15722
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: FR 00/06883
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: FR 00/12008
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatenIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 8101
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: JM109 / pMAL - Hs Cdc25C strain
US-10-149-472-5

Query Match 82.0%; Score 2705.8; DB 17; Length 8101;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGTTGCAAAACCTTTCCGGTATGGCATGATAGCGCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGTTGCAAAACCTTTCCGGTATGGCATGATAGCGCCGGAAGAGA 60

QY 61 GTCATTCAGGGTGGTGAATGTAACAGTAAAGTATACGATGTCGAGATGATGCGG 120
Db 61 GTCATTCAGGGTGGTGAATGTAACAGTAAAGTATACGATGTCGAGATGATGCGG 120

QY 121 GTGCTCTTATCAGACCGTTTCCCGCGTGGTGAACAGCCAGCCAGTTCCTGCGAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCCGCGTGGTGAACAGCCAGCCAGTTCCTGCGAAA 180

QY 181 CGCGGGAAGAGTGGAGCGCGATGCGGAGCTGAATTCATTCCTCCAAACCGCGTGGC 240
Db 181 CGCGGGAAGAGTGGAGCGCGATGCGGAGCTGAATTCATTCCTCCAAACCGCGTGGC 240

QY 241 AACAACTGGCGGGAACAGTCTGTTGCTGATGTCGCTGTCACCTCCAGTCTGCGCCCTGC 300
Db 241 AACAACTGGCGGGAACAGTCTGTTGCTGATGTCGCTGTCACCTCCAGTCTGCGCCCTGC 300

QY 301 ACGCGCGTGGCAAAATGTCGCGCGATTAATCTCGCGCGATCAACTCGGTCGCAAGC 360
Db 301 ACGCGCGTGGCAAAATGTCGCGCGATTAATCTCGCGCGATCAACTCGGTCGCAAGC 360

QY 361 TGGTGGTTCGATGGTGAAGCGCGCTGGAAGCTGTAAGCGCGCGTGGCAATC 420
Db 361 TGGTGGTTCGATGGTGAAGCGCGCTGGAAGCTGTAAGCGCGCGTGGCAATC 420

QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTACGCTGGATGACGAGTGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTACGCTGGATGACGAGTGCCA 480

QY 481 TTGCTGGAGAGTGGCTGCACTAAATGTTCCGGCTTATTTCTTGATGTTCTTGACAGA 540
Db 481 TTGCTGGAGAGTGGCTGCACTAAATGTTCCGGCTTATTTCTTGATGTTCTTGACAGA 540

Db 481 TTGCTGGAGAGTGGCTGCACTAAATGTTCCGGCTTATTTCTTGATGTTCTTGACAGA 540
QY 541 CACCATCAACAGTATTTATTTTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCATCAACAGTATTTATTTTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
QY 601 TGGTTCGATTTGGGTACACGCAAAATCGCTGTTAGCGGGCCCAATTAAATTTCTCTCGG 660
Db 601 TGGTTCGATTTGGGTACACGCAAAATCGCTGTTAGCGGGCCCAATTAAATTTCTCTCGG 660
QY 661 CGGCTCTGCGTCTGGCTGGGTCGATAAATATCTCACTCGCAATCAAATTCAGCGGATAG 720
Db 661 CGGCTCTGCGTCTGGCTGGGTCGATAAATATCTCACTCGCAATCAAATTCAGCGGATAG 720
QY 721 CGGAACGGGAAGCGACTGGAGTGGCCTGCTGGTTCGCGTTTCAACAAACCATGCAATTCGTA 780
Db 721 CGGAACGGGAAGCGACTGGAGTGGCCTGCTGGTTCGCGTTTCAACAAACCATGCAATTCGTA 780
QY 781 ATGAGGCGATCGTTTCCCATCTGCGATGCTGGTTCGCAACGATCAGATGGCGCTGGCGCAA 840
Db 781 ATGAGGCGATCGTTTCCCATCTGCGATGCTGGTTCGCAACGATCAGATGGCGCTGGCGCAA 840
QY 841 TGGCGGCATTCAGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 900
Db 841 TGGCGGCATTCAGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 900
QY 901 ACGATACCGGAAGACAGTCACTGTTATATCCGCGCTGCTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 901 ACGATACCGGAAGACAGTCACTGTTATATCCGCGCTGCTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
QY 961 GCCTGTGGGGCAAAACCGAGTGGACCGCTTGTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1080
Db 961 GCCTGTGGGGCAAAACCGAGTGGACCGCTTGTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGCTTCCCGCGTCTCACTGGTGAAGAAAACCAACCTGGCGCCCAATA 1140
Db 1081 CGCAAAACCGCTCTCCCGCGCTTCCCGCGTCTCACTGGTGAAGAAAACCAACCTGGCGCCCAATA 1140
QY 1141 CCGCATGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTGAGTTAGCTCACTCAATTAG 1200
Db 1141 CCGCATGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTGAGTTAGCTCACTCAATTAG 1200
QY 1201 GCACAAATCTCATGTTGACGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTGACGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGCGAGCCATCGGAAGCTGTGCTGATGGCTGCAAGGTGCTAAATCACTGCAATAATTCG 1320
Db 1261 TCAGGCGAGCCATCGGAAGCTGTGCTGATGGCTGCAAGGTGCTAAATCACTGCAATAATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGTCTGGATAATGTTTTTGGCGCGACATCAATACCGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTGGATAATGTTTTTGGCGCGACATCAATACCGTT 1380
QY 1381 CTGGCAAAATCTGGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAAAATCTGGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
QY 1441 ATTGTGAGCGGATAACAAATTTCAACAGGAAACAGCGCTCCGTTTGTGTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTTCAACAGGAAACAGCGCTCCGTTTGTGTTTTCACGA 1500
QY 1501 GCACCTTCACCAAGGACCATAGATTATGAAGCTGAAGAGGTAAATCTGGTAATCTCG 1560
Db 1501 GCACCTTCACCAAGGACCATAGATTATGAAGCTGAAGAGGTAAATCTGGTAATCTCG 1560
QY 1561 ATTAAACGGCGATAAAGGCTATAAACCGTCTCGCTGAGTAAAGAAATTCGAGAAGAT 1620
Db 1561 ATTAAACGGCGATAAAGGCTATAAACCGTCTCGCTGAGTAAAGAAATTCGAGAAGAT 1620


```
Qy 1621 ACCGGAATTAAGTACCGTTGAGCATCCGGATAAATCTGGAAGAGAAAAATCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTACCGTTGAGCATCCGGATAAATCTGGAAGAGAAAAATCCACAGGTT 1680

Qy 1681 GCGGCAATCGCGATGCGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAATCGCGATGCGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740

Qy 1741 GCTCAATCTGCGCTGTGGCTGAATATCAACCCCGGACAAAGCGTTCCAGGACAAGCTGAT 1800
Db 1741 GCTCAATCTGCGCTGTGGCTGAATATCAACCCCGGACAAAGCGTTCCAGGACAAGCTGAT 1800

Qy 1801 CCGTTTACCTGGATGCGGTAGCTTTACAAACGCAAGCTGATGCTTACCAGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGGTAGCTTTACAAACGCAAGCTGATGCTTACCAGATCGCTGTT 1860

Qy 1861 GAAGGTTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCCGCCAAACCTGGGAA 1920
Db 1861 GAAGGTTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCCGCCAAACCTGGGAA 1920

Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGTAAAGAGCGCGCTGATGTTCAAC 1980

Qy 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGTGACGGGGTTATGCGTTCAAG 2040

Qy 2041 TATGAARACGCAAGTACGACATTAAGACGTTGGCGCTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAARACGCAAGTACGACATTAAGACGTTGGCGCTGGATTAACGCTGGCGGAAAGCG 2100

Qy 2101 GGTCTGACCTTCTCTGTTGACCTGATTAACAAACACATGAAATGCAACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTAACAAACACATGAAATGCAACACCGATTAC 2160

Qy 2161 TCCATCGCAAGAGTGCCTTTAATAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
Db 2161 TCCATCGCAAGAGTGCCTTTAATAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220

Qy 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTTGTTGTTGTAACGGTACTCGCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTTGTTGTTGTAACGGTACTCGCGACCTTC 2280

Qy 2281 AAGGTCACACATCCAAACCGTTCTGTCGCTGCTGAGCGAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTCACACATCCAAACCGTTCTGTCGCTGCTGAGCGGAGGTATTAACGCCGCCAGT 2340

Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400

Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGTCCGTAGCGCTGAAGTCTTACGAGAAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGTCCGTAGCGCTGAAGTCTTACGAGAAAG 2460

Qy 2461 TTGGCGAAGATCCACGTTATGCGCCACACCTGGAAGAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGTTATGCGCCACACCTGGAAGAACGCCAGAAAGGTGAATCATG 2520

Qy 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCGTACTGCGGTGATCAACGCC 2580

Qy 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGTCTG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGTCTG 2640

Qy 2641 AACACAAACAACAATAACAATAACAACCAACCTCGGGATCGAGGAGGATTTTCAAAATTC 2700
Db 2641 AACACAAACAACAATAACAATAACAACCAACCTCGGGATCGAGGAGGATTTTCAAAATTC 2700
```

```
Qy 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709

RESULT 8
US-10-343-859-8
; Sequence 8, Application US/10343859
; Publication No. US20040110161A1
; GENERAL INFORMATION:
; APPLICANT: Nanogen Recognomics GMBH
; TITLE OF INVENTION: Method for detecting mutations in
; FILE OF INVENTION: nucleotide sequences
; FILE REFERENCE: 612,406-033
; CURRENT APPLICATION NUMBER: US/10/343,859
; CURRENT FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: PCT/EP01/08127
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 10038237.1
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 6648
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: plasmid
; OTHER INFORMATION: pMAL-c2x
US-10-343-859-8
```

Query Match 81.8%; Score 2701; DB 19; Length 6648;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2704; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```
Qy 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATGATAGCCCGGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATGATAGCCCGGGAAGAGA 60

Qy 61 GTCAATTCAAGGTTGTAATGTGAAACCAAGTAAAGCTTATACGATGTCGACAGTATGCGG 120
Db 61 GTCAATTCAAGGTTGTAATGTGAAACCAAGTAAAGCTTATACGATGTCGACAGTATGCGG 120

Qy 121 GTGTCCTTATCAGACCGTTTCCCGCGTGTGTAACAGCCAGCCAGCCACCTTTCTGCGAAA 180
Db 121 GTGTCCTTATCAGACCGTTTCCCGCGTGTGTAACAGCCAGCCAGCCACCTTTCTGCGAAA 180

Qy 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240

Qy 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300
Db 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300

Qy 301 ACGCCCGCTCGCAAAATTTGTCGCGCGCATTTAAATCTCGCGCCGATCAACTGGGTGCGCAGC 360
Db 301 ACGCCCGCTCGCAAAATTTGTCGCGCGCATTTAAATCTCGCGCCGATCAACTGGGTGCGCAGC 360

Qy 361 TGGTGGTGTGATGGTAGAACGAGCGCGTCGAAGCCCTGTAAGCGCGGTGCAACAATC 420
Db 361 TGGTGGTGTGATGGTAGAACGAGCGCGTCGAAGCCCTGTAAGCGCGGTGCAACAATC 420

Qy 421 TTCTCGCGCAACCGCTGATGGGCTGATTAATTAATCTATCCCTGATGATCAACAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTGATGGGCTGATTAATTAATCTATCCCTGATGATCAACAGGATGCCA 480

Qy 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 540

Qy 541 CACCCATCAACAGTATTTATTTTCTCCCATGAAGCGGTACGGGACTGGGGCTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTATTTTCTCCCATGAAGCGGTACGGGACTGGGGCTGGAGCATC 600
```

Db 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGCGCGTGGAGCATC 600
Qy 601 TGGTCGCAATTGGGTCAACAGCAAAATCGCGCTGTTAGCGGCCCAATTAAGTTCTGTCGCG 660
Db 601 TGGTCGCAATTGGGTCAACAGCAAAATCGCGCTGTTAGCGGCCCAATTAAGTTCTGTCGCG 660
Qy 661 CCGGCTCTGGGCTGGCTGGCTGGCAATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CCGGCTCTGGGCTGGCTGGCTGGCAATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CCGAAACGGGAAGCGCACATGGAGTGCATGTCGGTTTTCACAAACCAATGCAATGCTGA 780
Db 721 CCGAAACGGGAAGCGCACATGGAGTGCATGTCGGTTTTCACAAACCAATGCAATGCTGA 780
Qy 781 ATGAGGGCATCGTTCCTCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGTGGCGCAA 840
Db 781 ATGAGGGCATCGTTCCTCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGTGGCGCAA 840
Qy 841 TCGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTATGGGATACG 900
Db 841 TCGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTATGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAACCCACCATCAAAACAGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAACCCACCATCAAAACAGATTTTC 960
Qy 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGCGCCCTCTCACTGGTGAAGAAACACACCTGCGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGCGCCCTCTCACTGGTGAAGAAACACACCTGCGCGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGCGTGTGGCGATTCAATTAATGCAGTGGCAGCAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGCGTGTGGCGATTCAATTAATGCAGTGGCAGCAGGTTT 1140
Qy 1141 CCGGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTCTAGCTCACTATTAG 1200
Db 1141 CCGGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTCTAGCTCACTATTAG 1200
Qy 1201 GCACAATTCATGTTTGACAGCTTATCATCGATGCAAGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTCATGTTTGACAGCTTATCATCGATGCAAGTGCACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGATGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGATGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1320
Qy 1321 TGTGCTCAAGGCGCATCCCGTTCTGGATAATGTTTTCGCGCGGACATCAACCGTTT 1380
Db 1321 TGTGCTCAAGGCGCATCCCGTTCTGGATAATGTTTTCGCGCGGACATCAACCGTTT 1380
Qy 1381 CTGGCAAAATTTCTGAAATGAGTGTGACAAATTAATCATCGGCTCGTAAATGTGTGGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGTGTGACAAATTAATCATCGGCTCGTAAATGTGTGGA 1440
Qy 1441 ATTGTAGCGGATNAACAATTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTACGA 1500
Db 1441 ATTGTAGCGGATNAACAATTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTACGA 1500
Qy 1501 GCATCTCACCACCAAGGACCATAGATTATGAAACCTGAAGGTAACCTGGTAATCTGG 1560
Db 1501 GCATCTCACCACCAAGGACCATAGATTATGAAACCTGAAGGTAACCTGGTAATCTGG 1560
Qy 1561 ATTAACGGGATTAAGGCTATACGGTCTCGCTGAAGTCCGGTAAGAAATTCGAGAAGAT 1620
Db 1561 ATTAACGGGATTAAGGCTATACGGTCTCGCTGAAGTCCGGTAAGAAATTCGAGAAGAT 1620
Qy 1621 ACCGGAATTAAGTACCGTTGAGCATCCGGATAAATCTGGAAAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTACCGTTGAGCATCCGGATAAATCTGGAAAGAAATTCACAGGTT 1680

Qy 1581 CGGCAAACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACACGACCGCTTTGGTGGCTAC 1740
Db 1581 CGGCAAACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGGACAAAGGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGGACAAAGGTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACGGCAAGCTGATGTTCTTACCGGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACGGCAAGCTGATGTTCTTACCGGATCGCTGTT 1860
Qy 1861 GAGCGTTATCGCTGATTTATTAACAAGATCTGCTGCCGAACCGGCCAAAACCTGGGAA 1920
Db 1861 GAGCGTTATCGCTGATTTATTAACAAGATCTGCTGCCGAACCGGCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTACCTGCGCGCTGATTTGCTGTGAACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTGCGCGCTGATTTGCTGTGAACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAAACGGCAAGTAGACATTAAAGAGCTGGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAAACGGCAAGTAGACATTAAAGAGCTGGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGTTGACCTGATTTAAAAACAAACACATGAATGCAGACACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTTAAAAACAAACACATGAATGCAGACACCGGATTAC 2160
Qy 2161 TCCATGCGAGAACCTGCTTTTAAAGGGGAAACAGCGATGACCATCAACCGGCCGCTGG 2220
Db 2161 TCCATGCGAGAACCTGCTTTTAAAGGGGAAACAGCGATGACCATCAACCGGCCGCTGG 2220
Qy 2221 GCATGCTCAACATCGACACACAGCAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACACAGCAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
Qy 2281 AAGGTCACACCTTCAAAACCGTTTGGCTGTGAGCGCAGGTATTAACCGCCGCACT 2340
Db 2281 AAGGTCACACCTTCAAAACCGTTTGGCTGTGAGCGCAGGTATTAACCGCCGCACT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGTAGCGCTGAAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGTAGCGCTGAAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGGCGAAGATCCAGTATTCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCAGTATTCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGAGATGTCGCTTTCTGGTATGCGTGGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCCGAGATGTCGCTTTCTGGTATGCGTGGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Qy 2581 GCCAGCGTGTGAGTGTGATGAAGCCTTGAAGAGCGCGAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTGTGAGTGTGATGAAGCCTTGAAGAGCGCGAGACTAAATTCGAGCTCG 2640
Qy 2641 AACACAAACAAATAAATAAACAACCTCGGGATCGAGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACACAAACAAATAAATAAACAACCTCGGGATCGAGGAAGGATTTTCAGAAATTC 2700
Qy 2701 GGATCTCTCT 2709
Db 2701 GGATCTCTCT 2709

RESULT 9									
US-10-343-859-9									
; Sequence 9, Application US/10343859									
; Publication No. US20040110161A1									
; GENERAL INFORMATION:									
; APPLICANT: Nanogen Recognomics GMBH									
; TITLE OF INVENTION: Method for detecting mutations in									
; FILE OF INVENTION: nucleotide sequences									
; FILE REFERENCE: 612.406-033									
; CURRENT APPLICATION NUMBER: US/10/343,859									
; PRIOR FILING DATE: 2003-11-24									
; PRIOR APPLICATION NUMBER: PCT/EP01/08127									
; PRIOR FILING DATE: 2001-07-13									
; PRIOR APPLICATION NUMBER: 10038237.1									
; PRIOR FILING DATE: 2000-08-04									
; NUMBER OF SEQ ID NOS: 52									
; SOFTWARE: PatentIn Ver. 2.1									
; SEQ ID NO 9									
; LENGTH: 9191									
; TYPE: DNA									
; ORGANISM: Artificial sequence									
; FEATURE:									
; OTHER INFORMATION: Description of the artificial sequence: pMALc2x-muts									
; OTHER INFORMATION: Plasmid									
US-10-343-859-9									
Query Match 81.8%; Score 2698; DB 19; Length 9191;									
Best Local Similarity 99.8%; Pred. No. 0;									
Matches 2701; Conservative 0; Mismatches 5; Indels 0; Gaps 0;									
QY	1	CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGAGCCCGGAAGAGA	60						
DB	1	CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGAGCCCGGAAGAGA	60						
QY	61	GTCAATTACGGGTGGTGAATGTGAACAGTAACTATACCATGTCGCAGAGTATCGG	120						
DB	61	GTCAATTACGGGTGGTGAATGTGAACAGTAACTATACCATGTCGCAGAGTATCGG	120						
QY	121	GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACAGGCCAGCCACGTTTCTGCGAATA	180						
DB	121	GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACAGGCCAGCCACGTTTCTGCGAATA	180						
QY	181	CCGGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTTACATTTCCAAACCGGTGGCAC	240						
DB	181	CCGGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTTACATTTCCAAACCGGTGGCAC	240						
QY	241	AACAATCGCGGGCAACAGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300						
DB	241	AACAATCGCGGGCAACAGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300						
QY	301	ACGCGCGTTCGCAAAATTTGCGCGGATTAATACTTCGCGCGATCAACTGGGTGGCAGCG	360						
DB	301	ACGCGCGTTCGCAAAATTTGCGCGGATTAATACTTCGCGCGATCAACTGGGTGGCAGCG	360						
QY	361	TGGTGTGTCGATGTAGAACGAAGCGCGTGAAGCCCTGTAAAGCGCGGTGCAATC	420						
DB	361	TGGTGTGTCGATGTAGAACGAAGCGCGTGAAGCCCTGTAAAGCGCGGTGCAATC	420						
QY	421	TTCTCGCGCAACCGTCAGTGGCTGATCAATTAATTCGCTGGATGATGACAGGATGCCA	480						
DB	421	TTCTCGCGCAACCGTCAGTGGCTGATCAATTAATTCGCTGGATGATGACAGGATGCCA	480						
QY	481	TTGCTGTGAAGCTGCCTCACTAATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA	540						
DB	481	TTGCTGTGAAGCTGCCTCACTAATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA	540						
QY	541	CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGGGACTGGCGGTGGAGCATC	600						
DB	541	CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGGGACTGGCGGTGGAGCATC	600						
QY	601	TGGTGTGATGGTCAACGAAATCGCGCTGTAGCGGGCCCATTTAAGTTCTGTCTCGG	660						
DB									

DB	601	TGGTGGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG	660
QY	661	CGCGTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	661	CGCGTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTCAACAAACCAATGCAATGCTGA	780
DB	721	CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTCAACAAACCAATGCAATGCTGA	780
QY	781	ATGAGGGCATCGTTCCTCCACTCGCATGCTGTTGCCAAGATCAGATGCGCTGGCGCGAA	840
DB	781	ATGAGGGCATCGTTCCTCCACTCGCATGCTGTTGCCAAGATCAGATGCGCTGGCGCGAA	840
QY	841	TGCGCGCATTAACCGAGTCCCGGCTGCGGTGGTGGGATATCTCGGTAGTGGGATAG	900
DB	841	TGCGCGCATTAACCGAGTCCCGGCTGCGGTGGTGGGATATCTCGGTAGTGGGATAG	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCCACCATCAACAGGATTTTC	960
DB	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCCACCATCAACAGGATTTTC	960
QY	961	GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCGCCAGCGGTGA	1020
DB	961	GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCGCCAGCGGTGA	1020
QY	1021	AGGCAATCAGCTGTGTCCTCTCATGTTGAAAGAAAAACCAACCTGGCGCCCAATA	1080
DB	1021	AGGCAATCAGCTGTGTCCTCTCATGTTGAAAGAAAAACCAACCTGGCGCCCAATA	1080
QY	1081	CGAATACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGTGGCAGCAGAGTTT	1140
DB	1081	CGAATACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGTGGCAGCAGAGTTT	1140
QY	1141	CCGACTCGAAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTTAGTCTACTCATTTAG	1200
DB	1141	CCGACTCGAAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTTAGTCTACTCATTTAG	1200
QY	1201	GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGTGCACCAATGCTTCTGGCG	1260
DB	1201	GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGTGCACCAATGCTTCTGGCG	1260
QY	1261	TCAGCAGCCATCGGAAGCTGTGATGCTGTGAGGTGCTGAATTCATCTGATTAATTCG	1320
DB	1261	TCAGCAGCCATCGGAAGCTGTGATGCTGTGAGGTGCTGAATTCATCTGATTAATTCG	1320
QY	1321	TGTGCTCAAGCGCACCTCCCGTTCTGGATTAATGTTTTTGGCGCGACATCATACGGTT	1380
DB	1321	TGTGCTCAAGCGCACCTCCCGTTCTGGATTAATGTTTTTGGCGCGACATCATACGGTT	1380
QY	1381	CTGGCAAAATTTCTGAAAATGAGCTGTTGACAATTAATCATCGGCTCGTATATGTGGA	1440
DB	1381	CTGGCAAAATTTCTGAAAATGAGCTGTTGACAATTAATCATCGGCTCGTATATGTGGA	1440
QY	1441	ATTGTGAGCGGATTAACAATTTTCAACAGGAACAGCCAGTCCGTTAGTGTGTTTCA	1500
DB	1441	ATTGTGAGCGGATTAACAATTTTCAACAGGAACAGCCAGTCCGTTAGTGTGTTTCA	1500
QY	1501	GCACCTTCAACCAAGGACCATAGATATGAAAATGAAAGAGGTAATACTGTAATCTGG	1560
DB	1501	GCACCTTCAACCAAGGACCATAGATATGAAAATGAAAGAGGTAATACTGTAATCTGG	1560
QY	1561	ATTAACGGCGATTAAGGCTATAACCGCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
DB	1561	ATTAACGGCGATTAAGGCTATAACCGCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCGGATAAATGGAAGAGAAATTCGCCAGGTT	1680
DB	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCGGATAAATGGAAGAGAAATTCGCCAGGTT	1680
QY	1681	GGGCAATCGGCGATGGCCCTGACATTTCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
DB	1681	GGGCAATCGGCGATGGCCCTGACATTTCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740

```
QY 1741 GCTCAATCTGGCTGTGGCTGAATCATCCCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAATCATCCCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCCGTAGCTTTACAAAGCAAGCTGATTTGCTTACCCGATCCTGTT 1860
Db 1801 CCGTTTACCTGGATGCCGTAGCTTTACAAAGCAAGCTGATTTGCTTACCCGATCCTGTT 1860
QY 1861 GAAGCGTTATCCTGCTGATTTATAACAAAGATCTGTCGCGAACCCCGCCAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCCTGCTGATTTATAACAAAGATCTGTCGCGAACCCCGCCAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCAAGGTAAAGAGCGCGCTGATTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCAAGGTAAAGAGCGCGCTGATTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACAGTGGSCGTGGATTAAGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACAGTGGSCGTGGATTAAGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGCTGATTAAGCACTGATTAATAAACAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGATTAAGCACTGATTAATAAACAACACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTAAATAAGCGCAACACAGCGATGACCATCAACGGCGCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTAAATAAGCGCAACACAGCGATGACCATCAACGGCGCGTGG 2220
QY 2221 GCATGGTCAACATCGACACAGCAAAAGTGAATTAATGTTGTAACGTTACTCCGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACAGCAAAAGTGAATTAATGTTGTAACGTTACTCCGACCTTC 2280
QY 2281 AAGGTCAACCATCAACCGTTGCTGGCGTGTGAGCGCAGGTATTAACCGCGCCAGT 2340
Db 2281 AAGGTCAACCATCAACCGTTGCTGGCGTGTGAGCGCAGGTATTAACCGCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACCTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACCTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGTTAAATAAGCAAAACCGCTGGGTGCCGTAGCGCTGAAAGTCTTACGAGGAAAG 2460
Db 2401 GAAGCGTTAAATAAGCAAAACCGCTGGGTGCCGTAGCGCTGAAAGTCTTACGAGGAAAG 2460
QY 2461 TTGGCGAAAGATCCACGTATTTGCCGCCACCATGGAAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTATTTGCCGCCACCATGGAAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATCCGTTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATCCGTTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTGTGAGATGTCGATGAAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTGTGAGATGTCGATGAAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACCAACACACATACATTAACAAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACCAACACACATTAACAAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
QY 2701 GGATCC 2706
Db 2701 GGATCC 2706
```

RESULT 10
US-10-263-153-40
Sequence 40, Application US/10263153

```
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Gineburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263.153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 7112
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3189)
; OTHER INFORMATION: pMBP-c2K-Toxop30del11 (52-214aa)
; US-10-263-153-40
```

```
Query Match 81.5% Score 2690.4; DB 18; Length 7112;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGTATGGCATGATGCGCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGTATGGCATGATGCGCCGGAAGAGA 60
QY 61 GTCAATTCAGGGTGGTGAATGTGAACCCAGTAACGTTATACGATGTCGAGAGTAGCCG 120
Db 61 GTCAATTCAGGGTGGTGAATGTGAACCCAGTAACGTTATACGATGTCGAGAGTAGCCG 120
QY 121 GTCTCTTTATCAGACCGTTTCCCGGCTGGTGAACAGGCGCAGCCAGCTTCTTCCGAAA 180
Db 121 GTCTCTTTATCAGACCGTTTCCCGGCTGGTGAACAGGCGCAGCCAGCTTCTTCCGAAA 180
QY 181 CCGCGGAAAAGTGGAAAGCGGCGATGGCGAGCTGAATTAATTCCTCCCAACCGCGTGGC 240
Db 181 CCGCGGAAAAGTGGAAAGCGGCGATGGCGAGCTGAATTAATTCCTCCCAACCGCGTGGC 240
QY 241 AACACTGGCGGCAACAGTCTGCTGATTTGGCGCTTGCACCTCCAGTCTGGCCCTGC 300
Db 241 AACACTGGCGGCAACAGTCTGCTGATTTGGCGCTTGCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCCCGTGC AAAATTGTGCGCGCGATTAATAATCTCGCGCCGATCAACTGGGTGCCAG 360
Db 301 ACGCCCGTGC AAAATTGTGCGCGCGATTAATAATCTCGCGCCGATCAACTGGGTGCCAG 360
QY 361 TGGTGTGTGCGATGGTAGAACGAGCGGCTGAAAGCTGTAAAGCGGCGGTGCACAATC 420
Db 361 TGGTGTGTGCGATGGTAGAACGAGCGGCTGAAAGCTGTAAAGCGGCGGTGCACAATC 420
QY 421 TTCTCCGCAACCGCTCAGTGGCTGATCAATTAATTCCTCGCTGGATGACCGAGTGGCA 480
Db 421 TTCTCCGCAACCGCTCAGTGGCTGATCAATTAATTCCTCGCTGGATGACCGAGTGGCA 480
QY 481 TTGCTGTGAAGCTGCTGCACCTTAATTTCCGCGTTATTTCTTGATGTCTCTGACCAGA 540
Db 481 TTGCTGTGAAGCTGCTGCACCTTAATTTCCGCGTTATTTCTTGATGTCTCTGACCAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTAGCGGACTGGGCGGTGGAGCAT 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTAGCGGACTGGGCGGTGGAGCAT 600
QY 601 TGGTGTGATTTGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTCTGTCTCG 660
Db 601 TGGTGTGATTTGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTCTGTCTCG 660
QY 661 CCGGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
```


/ TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
/ TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
/ FILE REFERENCE: 6984.US.01
/ CURRENT APPLICATION NUMBER: US/10/263,153
/ CURRENT FILING DATE: 2002-10-22
/ NUMBER OF SEQ ID NOS: 74
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 30
/ LENGTH: 7259
/ TYPE: DNA
/ ORGANISM: Toxoplasma gondii
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1528)...(3336)
/ OTHER INFORMATION: pMBP-c2X-Toxop30del14del18 (83-294aa)
US-10-263-153-30

Query Match 81.5%; Score 2690.4; DB 18; Length 7259;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGCAGAACTTTCCGGGTATGGCATGATAGCGCCGGAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAGAACTTTCCGGGTATGGCATGATAGCGCCGGAGAGA 60
QY 61 GTCGAATTCAGGGTGGTGAATGTGAACCCAGTAACGTTATACGATGTTCGACAGTATGCCG 120
Db 61 GTCGAATTCAGGGTGGTGAATGTGAACCCAGTAACGTTATACGATGTTCGACAGTATGCCG 120
QY 121 GTGCTCTTATCAGACCGGTTTCCCGGTGGTGAACCGGCGAGCCAGCCACGTTTCTCGAAAA 180
Db 121 GTGCTCTTATCAGACCGGTTTCCCGGTGGTGAACCGGCGAGCCAGCCACGTTTCTCGAAAA 180
QY 181 CCGGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTAATACATCCCAACCGCGTGCAC 240
Db 181 CCGGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTAATCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAAAACAGTGGTGTGATTTGGCGTTTGGCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGCAAAACAGTGGTGTGATTTGGCGTTTGGCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGGCGCTCGCAATTTGTCGGCGGATTAATCTCGCGCGGATCACTGGTGGCGAGCG 360
Db 301 ACGGCGCTCGCAATTTGTCGGCGGATTAATCTCGCGCGGATCACTGGTGGCGAGCG 360
QY 361 TGGTGGTTCGATGGTGAACGAAGCGCGCTCGAAGCGCTGAAGCGCTGAAGCGCGGTGCAATC 420
Db 361 TGGTGGTTCGATGGTGAACGAAGCGCGCTCGAAGCGCTGAAGCGCGGTGCAATC 420
QY 421 TTCTCGGCAACGGCTCAGTGGGCTGATCAATTAATCTATCGGCGGATCACTGGATGCCA 480
Db 421 TTCTCGGCAACGGCTCAGTGGGCTGATCAATTAATCTATCGGCGGATCACTGGATGCCA 480
QY 481 TTGCTGTGGAGCTGCCTGCACTAATGTTCCGGGTTATTTCTTGATGTCTCTGACCAGA 540
Db 481 TTGCTGTGGAGCTGCCTGCACTAATGTTCCGGGTTATTTCTTGATGTCTCTGACCAGA 540
QY 541 CACCCATCAACAGTATTTTCTCCATGAAGACGGTACGGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTTCTCCATGAAGACGGTACGGACTGGGCGTGGAGCATC 600
QY 601 TGGTTCGATTTGGGTACACAGCAATTCGGCTGTAGGGGCCCATTTAAGTTCGTCTCGG 660
Db 601 TGGTTCGATTTGGGTACACAGCAATTCGGCTGTAGGGGCCCATTTAAGTTCGTCTCGG 660
QY 661 CGCGTCTCGCTTGGCTGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATG 720
Db 661 CGCGTCTCGCTTGGCTGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATG 720
QY 721 CGGAACGGGAAGCGCATGGAGTGCCATGTCGGTTTCAACAAACCATCAATGCTGA 780
Db 721 CGGAACGGGAAGCGCATGGAGTGCCATGTCGGTTTCAACAAACCATCAATGCTGA 780

QY 781 ATGAGGCGATCGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGCGCGCTGGCGCAA 840
Db 781 ATGAGGCGATCGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGCGCGCTGGCGCAA 840
QY 841 TGGCGGCATTAACGAGTTCGGGCTGCGGCTGGTGGGATATCTCGTGTAGTGGATACG 900
Db 841 TGGCGGCATTAACGAGTTCGGGCTGCGGCTGGTGGGATATCTCGTGTAGTGGATACG 900
QY 901 ACGATACGGAAGACAGCTCATGTTATATCCGCGGTTAAACACCATCAAAACAGATTTTC 960
Db 901 ACGATACGGAAGACAGCTCATGTTATATCCGCGGTTAAACACCATCAAAACAGATTTTC 960
QY 961 GCCTGCTGGGGCAACACGAGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGCTGGGGCAACACGAGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAAAAACACCCCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAAAAACACCCCTGGCGCCCAATA 1080
QY 1081 CGMAACCGGCTCTCCCGCGGTTGGCCGATTCATTAATGACGCTGGCAGCAGAGTTT 1140
Db 1081 CGMAACCGGCTCTCCCGCGGTTGGCCGATTCATTAATGACGCTGGCAGCAGAGTTT 1140
QY 1141 CCGACTGGAAAGCGGCGAGTGGCAACGCAATTAATGAGTTAGCTCATTCTAG 1200
Db 1141 CCGACTGGAAAGCGGCGAGTGGCAACGCAATTAATGAGTTAGCTCATTCTAG 1200
QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACGACCATCGGAAGCTGGTATGCTGTGAGGCTGTAATCACTGCATAAATTCG 1320
Db 1261 TCAGGACGACCATCGGAAGCTGGTATGCTGTGAGGCTGTAATCACTGCATAAATTCG 1320
QY 1321 TGTGCTCAAGCGCGACTCCCGCTTCTGGATAATGTTTTTGGCGCGACATCAAAACGGTT 1380
Db 1321 TGTGCTCAAGCGCGACTCCCGCTTCTGGATAATGTTTTTGGCGCGACATCAAAACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
QY 1441 ATTGTGAGCGGATAACAAATTTACACAGGAAACAGCGAGTTCGGTTAGTGTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTTACACAGGAAACAGCGAGTTCGGTTAGTGTTCACGA 1500
QY 1501 GCACTTCCACCAACAGGACCATAGATTATGAAAACTGAAGAAAGGTAAACTGGTAAATCTGG 1560
Db 1501 GCACTTCCACCAACAGGACCATAGATTATGAAAACTGAAGAAAGGTAAACTGGTAAATCTGG 1560
QY 1561 ATTAAACGGGATTAAGGCTTAACCGTCTCGCTGAAAGTGGTAAAGAAATTCGAAAGAT 1620
Db 1561 ATTAAACGGGATTAAGGCTTAACCGTCTCGCTGAAAGTGGTAAAGAAATTCGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACGCTGAGCATCCGATTAACCTCGAAGAGAAATTCGACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGCTGAGCATCCGATTAACCTCGAAGAGAAATTCGACAGGTT 1680
QY 1681 GGGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACAGCGCTTGGTGGCTAC 1740
Db 1681 GGGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACAGCGCTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCACCCGGCAAAAGCGTTCCAGCAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATCACCCGGCAAAAGCGTTCCAGCAAGCTGTAT 1800
QY 1801 CGGTTTACCTGGGATGCCGATGCTTACACGCAAGCTGATTGCTTACCCGATCGCTGT 1860
Db 1801 CGGTTTACCTGGGATGCCGATGCTTACACGCAAGCTGATTGCTTACCCGATCGCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCGCAACCCGCGCAAAACCTGGGAA 1920


```
Db 1861 GAAGCGTTATCGCTGATTTTAAACAAGATGCTGCGGAACCCGCCAAAAACCTGGGAA 1920
Qy 1921 GAGATCCGGCGCTCGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGGCGCTCGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTCGGCGCTGATGCTGTGACGCGGGGTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTCGGCGCTGATGCTGTGACGCGGGGTATGCGTTCAAG 2040
Qy 2041 TATGAARACGCACTGACACATTAAGAAGCGTGGCGTGGATAAGCTGGCGGAAGCG 2100
Db 2041 TATGAARACGCACTGACACATTAAGAAGCGTGGCGTGGATAAGCTGGCGGAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGTGTGACCTGATTAATAAACAACACATGAAATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTGTGACCTGATTAATAAACAACACATGAAATGACAGACACCGATTAC 2160
Qy 2161 TCCATCGAGAAAGCTGCTTTAATAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGAGAAAGCTGCTTTAATAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCAACATCGACACCAAGAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACCAAGAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
Qy 2281 AAGGTCACCAATCAACACCGTTCGTTGGCGTGTGAGCGCAGGTATTAAACCGCCCACTG 2340
Db 2281 AAGGTCACCAATCAACACCGTTCGTTGGCGTGTGAGCGCAGGTATTAAACCGCCCACTG 2340
Qy 2341 CCGAAACAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGACTGATGAAGCTCTG 2400
Db 2341 CCGAAACAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGACTGATGAAGCTCTG 2400
Qy 2401 GAAGCGGTATTAAGAACAACCGCTGGTGGCTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTATTAAGAACAACCGCTGGTGGCTAGCGCTGAAGTCTTACGAGGAAG 2460
Qy 2461 TTGGGAAAGATCCACGATTTGCGCGCCACATGGAAGAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGGAAAGATCCACGATTTGCGCGCCACATGGAAGAACGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAATATCCCGCAGATGTCGCTTTCTGTTGATGCGGTGATGCGGTGATCAACGCC 2580
Db 2521 CCGAATATCCCGCAGATGTCGCTTTCTGTTGATGCGGTGATGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACACTGTCGATGAGCCCTGAAAGAGCGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACACTGTCGATGAGCCCTGAAAGAGCGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACAAACAACAATAACAATAACAACAACCTCGGATCGAGGAAAGGATTTCAAGATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACAACCTCGGATCGAGGAAAGGATTTCAAGATTC 2700
```

RESULT 12

```
US-10-263-153-35
; Sequence 35, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.O1
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 35
; LENGTH: 7322
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3399)
; OTHER INFORMATION: pMBP-c2X-Toxop30del10 (52-284aa)
US-10-263-153-35
```

Query Match

```
Best Local Similarity 81.5%; Score 2690.4; DB 18; Length 7322;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
Qy 1 CCGACACCATCGAATGGTGTGCAAAACCTTTTCGCGGTATGCGCATGATAGAGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGTGCAAAACCTTTTCGCGGTATGCGCATGATAGAGCCCGGAAGAGA 60
Qy 61 GTCAATTCCAGGGTGGTGAATGTGAACCCAGTAAAGTTATACGATGTGCGAGATATGCGG 120
Db 61 GTCAATTCCAGGGTGGTGAATGTGAACCCAGTAAAGTTATACGATGTGCGAGATATGCGG 120
Qy 121 GTGTCCTTTATCAGACCGCTTTCCCGCGTGTGTAACCCAGCCAGCCAGCTTTCTCGGAAAA 180
Db 121 GTGTCCTTTATCAGACCGCTTTCCCGCGTGTGTAACCCAGCCAGCCAGCTTTCTCGGAAAA 180
Qy 181 CGCGGAAAAAGTGAAGCGCGCATGCGCGAGCTGAATTACATTTCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGCGCATGCGCGAGCTGAATTACATTTCCCAACCGCGTGGCAC 240
Qy 241 AACAACTGGCGGGCAACAGTCTGTTGTTGTTGGGTTGGCCACTTCCAGTCTGGGCCCTGC 300
Db 241 AACAACTGGCGGGCAACAGTCTGTTGTTGTTGGGTTGGCCACTTCCAGTCTGGGCCCTGC 300
Qy 301 ACGCGCGTTCGCAAAATTTCTCGCGCGATTAAATCTCGCGCGATCAACTGGGTGCGAGCG 360
Db 301 ACGCGCGTTCGCAAAATTTCTCGCGCGATTAAATCTCGCGCGATCAACTGGGTGCGAGCG 360
Qy 361 TGTGTTGTCGATGATAGAACGAGCGCGCTGCAAGCCCTGTAAGCGCGGTGCAACAATC 420
Db 361 TGTGTTGTCGATGATAGAACGAGCGCGCTGCAAGCCCTGTAAGCGCGGTGCAACAATC 420
Qy 421 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATCTCGGTGATGATGATGATGATGATG 480
Db 421 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATCTCGGTGATGATGATGATGATGATG 480
Qy 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTATTTCTTCAATGTTCTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTATTTCTTCAATGTTCTCTGACCCAGA 540
Qy 541 CACCCATCAACAGATATTATTTTCTCCCATGAAGACGCTACGCGACTGGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGATATTATTTTCTCCCATGAAGACGCTACGCGACTGGGCGGTGGAGCATC 600
Qy 601 TGTGTCGATTTGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCTCGG 660
Db 601 TGTGTCGATTTGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCTCGG 660
Qy 661 CGCGTCTGGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CCGAAACGGGAAGCGCATGCGAGTGCATGTCGGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CCGAAACGGGAAGCGCATGCGAGTGCATGTCGGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATCTTCCCATCTGCGATGCTGGTTGCGCAACGATCAGATGGCGCTGGCGGCA 840
Db 781 ATGAGGGCATCTTCCCATCTGCGATGCTGGTTGCGCAACGATCAGATGGCGCTGGCGGCA 840
Qy 841 TCGCGGCCATTACCGAGTCCGGGCTGGCGGTGGTGGCGGATATCTCGTAGTGGGATAG 900
Db 841 TCGCGGCCATTACCGAGTCCGGGCTGGCGGTGGTGGCGGATATCTCGTAGTGGGATAG 900
```

```
QY 901 ACGATACCGAAGACAGCTCATGTTATATATCCCGCGTTAAACCACTCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATATCCCGCGTTAAACCACTCAAAACAGGATTTTC 960
QY 961 GCTGCTGGGGCAAAACACAGCGTGGACCGTTCCTGCACTCTCTCAGGCGCAGCGGTGA 1020
Db 961 GCTGCTGGGGCAAAACACAGCGTGGACCGTTCCTGCACTCTCTCAGGCGCAGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTTGCGCGCTCTCACTGCTGAAAGAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGCGCGCTCTCACTGCTGAAAGAAACCAACCTGGCGCCCAATA 1080
QY 1081 CCGAAACCGCTCTCCCGCGGTGCGCGATTCAATTAATGCACTGGCAGCAAGGTTT 1140
Db 1081 CCGAAACCGCTCTCCCGCGGTGCGCGATTCAATTAATGCACTGGCAGCAAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAAGTTAGCTCACTCAATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAAGTTAGCTCACTCAATTAG 1200
QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATGACTGCACGGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTTCTCATGTTTGACAGCTTATCATGACTGCACGGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACGCATCGGAAGCTGTGTTGCTGTGAGTGTGCACTGTAATCACTGCATAATTTCG 1320
Db 1261 TCAGGACGCATCGGAAGCTGTGTTGCTGTGAGTGTGCACTGTAATCACTGCATAATTTCG 1320
QY 1321 TGTGCTCAAGCGGCACCTCCCGTCTCGGATTAATGTTTTCGCGCGACATCAACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACCTCCCGTCTCGGATTAATGTTTTCGCGCGACATCAACGGTT 1380
QY 1381 CTGGCAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
Db 1381 CTGGCAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
QY 1441 ATTGTAGCGGATAACANTTTACACAGGAAACGCGTTCGTTTAGGTGTTTTCACGA 1500
Db 1441 ATTGTAGCGGATAACANTTTACACAGGAAACGCGTTCGTTTAGGTGTTTTCACGA 1500
QY 1501 GCATTTCAACCAACAGGACCATAGATTATGAAAACTGAAGAAGTAACTGGTAATCTGG 1560
Db 1501 GCATTTCAACCAACAGGACCATAGCATATGAAATCGAAGAGTAACTGGTAATCTGG 1560
QY 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAACTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAACTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCACCGTTGACATCCGGATAACTGGNAGAGAAATTCACACAGGTT 1680
Db 1621 ACCGGAATTAAGTCACCGTTGACATCCGGATAACTGGNAGAGAAATTCACACAGGTT 1680
QY 1681 GCGCAACTGCGGATGCGCCCTGACATTAATCTTCTGGGCACACGACCGCTTGGTGGCTAC 1740
Db 1681 GCGCAACTGCGGATGCGCCCTGACATTAATCTTCTGGGCACACGACCGCTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTTGGCTGAAATCAACCCCGACAAAGGTTTCCAGGCAAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTTGGCTGAAATCAACCCCGACAAAGGTTTCCAGGCAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGTACGTTACACGGCAAGCTGATTCGTTACCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGTACGTTACACGGCAAGCTGATTCGTTACCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCGGCCAAAACCTGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCGGCCAAAACCTGGAA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGACTGAAAGCGAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGACTGAAAGCGAAGGTAAGAGCGCGCTGATGTTCAAC 1980
```

```
QY 1981 CTGCAAGAACGCTACTTCACTGCGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACGCTACTTCACTGCGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTAGTACGACATTAAAGACGTTGGGCGTGGATACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTAGTACGACATTAAAGACGTTGGGCGTGGATACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAAAAAAACAACATGATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAAAAAAACAACATGATGACAGACACCGATTAC 2160
QY 2161 TCCATCCGAGAAGCTGCTTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCCGAGAAGCTGCTTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGCTCCAAACATCGACACCAAGTGAATTAATGTTGTAACGGTACTTGCAGACCTTC 2280
Db 2221 GCATGCTCCAAACATCGACACCAAGTGAATTAATGTTGTAACGGTACTTGCAGACCTTC 2280
QY 2281 AAGGTCCAACATCCCAACCGTTTGGTGGCTGCTGAGCGCAGGTATTAAACGCCGCCAGT 2340
Db 2281 AAGGTCCAACATCCCAACCGTTTGGTGGCTGCTGAGCGCAGGTATTAAACGCCGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGCGCAAAAGAGTTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGCGCAAAAGAGTTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTTAATAAGACAAACCGCTGGTGGTGGCTGAGCTGAAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTTAATAAGACAAACCGCTGGTGGTGGCTGAGCTGAAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGCGGAAAGATCCAGTATTGCGCGCACCATCGAAACGCCAGAAAGTGTAATCATG 2520
Db 2461 TTGCGGAAAGATCCAGTATTGCGCGCACCATCGAAACGCCAGAAAGTGTAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTGCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTGCTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGCTGAGACTGCTGATGAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGCTGAGACTGCTGATGAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAAATAACAAATAACAAATAACAAATAACAAATAACAAATAACAAATAAC 2700
Db 2641 AACAAACAAACAAATAACAAATAACAAATAACAAATAACAAATAACAAATAACAAATAAC 2700
```

RESULT 13

```
US-10-263-153-25
; Sequence 25, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; FILE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.O1
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 7352
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3429)
```


OTHER INFORMATION: pMBP-c2X-ToxoP30del14C (52-294aa)
us-10-263-153-25

Query Match 81.5%; Score 2690.4; DB 18; Length 7352;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy	1	CCGACACCATCGAATGGTGGCAAAACCTTTCCGGGTATGGCATGATAGGCCCGCGGAAGAGA	60
Db	1	CCGACACCATCGAATGGTGGCAAAACCTTTCCGGGTATGGCATGATAGGCCCGCGGAAGAGA	60
Qy	61	GTCAATTCCAGGCTGTGAATGTGAACACCAAGTAACTATACGATGTCGACAGATATGCCG	120
Db	61	GTCAATTCCAGGCTGTGAATGTGAACACCAAGTAACTATACGATGTCGACAGATATGCCG	120
Qy	121	GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACACGCGCAGCCACCGTTTCTCGAAAA	180
Db	121	GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACACGCGCAGCCACCGTTTCTCGAAAA	180
Qy	181	CGCGGAAAAAGTGAACCGCGCATGGCGGAGCTGAATTACATTCGCCAACCGCGTGGCAC	240
Db	181	CGCGGAAAAAGTGAACCGCGCATGGCGGAGCTGAATTACATTCGCCAACCGCGTGGCAC	240
Qy	241	AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGSCCCTGC	300
Db	241	AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGSCCCTGC	300
Qy	301	ACGGCCGCTCGCAAAATTTGCGCGCGATTAAATCTCGCGCGGATCAACTGGGTGCCAGCG	360
Db	301	ACGGCCGCTCGCAAAATTTGCGCGCGATTAAATCTCGCGCGGATCAACTGGGTGCCAGCG	360
Qy	361	TGGTGGTTCGATGGTAGAACGAAAGCGCGTGAAGCTGTAAAGCGCGGTGCAACAATC	420
Db	361	TGGTGGTTCGATGGTAGAACGAAAGCGCGTGAAGCTGTAAAGCGCGGTGCAACAATC	420
Qy	421	TTCTCGGCAACGGTCACTGGGCTGATCAATTAATCTCGCTGGATCAACGAGATGCCA	480
Db	421	TTCTCGGCAACGGTCACTGGGCTGATCAATTAATCTCGCTGGATCAACGAGATGCCA	480
Qy	481	TTGCTGTGGAAGCTGCCCTGCAATATGTTTCGGCGTTATTTCTTGATGCTCTGACACAGA	540
Db	481	TTGCTGTGGAAGCTGCCCTGCAATATGTTTCGGCGTTATTTCTTGATGCTCTGACACAGA	540
Qy	541	CACCCATCAACAGTATTTTCTCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC	600
Db	541	CACCCATCAACAGTATTTTCTCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC	600
Qy	601	TGGTGGATTTGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAAGTTCTGTCTCG	660
Db	601	TGGTGGATTTGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAAGTTCTGTCTCG	660
Qy	661	CGCGTCTGCGTCTGGCTGGCAATAATCTCACCTCGCAATCAAAATTCAGCCGATAG	720
Db	661	CGCGTCTGCGTCTGGCTGGCAATAATCTCACCTCGCAATCAAAATTCAGCCGATAG	720
Qy	721	CGGAAACGGAAGCGACTGGAGTGCATGTCCGGTTTTCAACAAACCATGCAATGCTGA	780
Db	721	CGGAAACGGAAGCGACTGGAGTGCATGTCCGGTTTTCAACAAACCATGCAATGCTGA	780
Qy	781	ATGAGGGATCGTTTCCCACTGGATGCTGGTTGCCAAGATCAGATGGCGCTGGCGGCA	840
Db	781	ATGAGGGATCGTTTCCCACTGGATGCTGGTTGCCAAGATCAGATGGCGCTGGCGGCA	840
Qy	841	TGCGGCGCATTTACAGATCGCGGCTGGCTGGTGGCGGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGGCGCATTTACAGATCGCGGCTGGCGTGGTGGCGGATATCTCGGTAGTGGGATACG	900
Qy	901	ACGATACGGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC	960
Db	901	ACGATACGGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC	960
Qy	961	GCCTGCTGGGCAAAACAGCGTGGACCGGTTGCTGCAACTCTCTCAGGCGCAGCGGTGA	1020
Db	961	GCCTGCTGGGCAAAACAGCGTGGACCGGTTGCTGCAACTCTCTCAGGCGCAGCGGTGA	1020

Db	961	GCCTGCTGGGCAAAACAGCGTGGACCGGTTGCTGCAACTCTCTCAGGCGCAGCGGTGA	1020
Qy	1021	AGGCAATCAGCTGTTGCCGTCTCCTGTGTGAAAGAAAAACCACTGCGGCCCAATA	1080
Db	1021	AGGCAATCAGCTGTTGCCGTCTCCTGTGTGAAAGAAAAACCACTGCGGCCCAATA	1080
Qy	1081	CGGAAACGGCTCTCCCGCGGTTGGCGGATTCATTAATGACGCTGGCAGCAGAGTTT	1140
Db	1081	CGGAAACGGCTCTCCCGCGGTTGGCGGATTCATTAATGACGCTGGCAGCAGAGTTT	1140
Qy	1141	CCGACTTGGAAAGCGGCGAGTGAGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG	1200
Db	1141	CCGACTTGGAAAGCGGCGAGTGAGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG	1200
Qy	1201	GCAAAATTCATGTTTGACAGCTTATCATCGATGCACTGCACTGCAACCAATGCTTCGGC	1260
Db	1201	GCAAAATTCATGTTTGACAGCTTATCATCGATGCACTGCACTGCAACCAATGCTTCGGC	1260
Qy	1261	TCAGGACGCAATCGGAAGCTGTGTATGCTGTGAGGTCTGTAATCACTGCATTAATTCG	1320
Db	1261	TCAGGACGCAATCGGAAGCTGTGTATGCTGTGAGGTCTGTAATCACTGCATTAATTCG	1320
Qy	1321	TGTCGCTCAAGCGCACCTCCCGTTCTGGATTAATGTTTTCGGCGGACATCAACAGGTT	1380
Db	1321	TGTCGCTCAAGCGCACCTCCCGTTCTGGATTAATGTTTTCGGCGGACATCAACAGGTT	1380
Qy	1381	CTGCAAAATTTCTGAAATGAGTGTGACAAATTAATCATCGGCTCGTAAATGTTGTTG	1440
Db	1381	CTGCAAAATTTCTGAAATGAGTGTGACAAATTAATCATCGGCTCGTAAATGTTGTTG	1440
Qy	1441	ATTGTAGCGGATTAACAAATTTTACACAGGAAACGCGTCCGTTAGGTGTTTTCACGA	1500
Db	1441	ATTGTAGCGGATTAACAAATTTTACACAGGAAACGCGTCCGTTAGGTGTTTTCACGA	1500
Qy	1501	GCCTTCCACCAAGGACCATAGATTATGAACTGAAAGAGGTAAATCGTAACTCTG	1560
Db	1501	GCCTTCCACCAAGGACCATAGATTATGAACTGAAAGAGGTAAATCGTAACTCTG	1560
Qy	1561	ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGTAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGTAAGAAATTCGAGAAAGAT	1620
Qy	1621	ACCGAATTAAGTCAACGGTTGAGGATCCGGATAAATCGGAAGAGAAATTCACAGGTT	1680
Db	1621	ACCGAATTAAGTCAACGGTTGAGGATCCGGATAAATCGGAAGAGAAATTCACAGGTT	1680
Qy	1681	CGGCACTGCGGATGCGCTGACATTAATCTTCTGGGACACGCGCTTTGGTGGCTAC	1740
Db	1681	CGGCACTGCGGATGCGCTGACATTAATCTTCTGGGACACGCGCTTTGGTGGCTAC	1740
Qy	1741	GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGGTTTCAGGACAAAGCTGAT	1800
Db	1741	GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGGTTTCAGGACAAAGCTGAT	1800
Qy	1801	CGGTTTACCTGGGATGCGGTACGTTAACAACGCAAGCTGATTGCTTACCGGATCGCTGT	1860
Db	1801	CGGTTTACCTGGGATGCGGTACGTTAACAACGCAAGCTGATTGCTTACCGGATCGCTGT	1860
Qy	1861	GAAGGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCAAAACCTGGGAA	1920
Db	1861	GAAGGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCAAAACCTGGGAA	1920
Qy	1921	GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980
Qy	1981	CTGCAAGAACCGTACTTTCACCTGCGCGTGAATGCTGCTGACGGGGGTTATGCGTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTTCACCTGCGCGTGAATGCTGCTGACGGGGGTTATGCGTTCAAG	2040
Qy	2041	TATGAAACGCAAGTACGACATTAAGACCTGGCGGTGATTAAGCTGGCGGCAAGCG	2100
Db	2041	TATGAAACGCAAGTACGACATTAAGACCTGGCGGTGATTAAGCTGGCGGCAAGCG	2100

```
QY 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAAAAACAACACATGAATGCACACCCGATTAC 2160
Db |||||
QY 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAAAAACAACACATGAATGCACACCCGATTAC 2160
Db |||||
QY 2161 TCCATCGCAGAGCTCCCTTTAATAAGCGCAACACAGCGATGACCATCAACCGCCCGTGG 2220
Db |||||
QY 2161 TCCATCGCAGAGCTCCCTTTAATAAGCGCAACACAGCGATGACCATCAACCGCCCGTGG 2220
Db |||||
QY 2221 GCATGGTCCAAATCGACACACAGCAAGTGAATTATGGTGAACGGTACTGCGGACCTTC 2280
Db |||||
QY 2221 GCATGGTCCAAATCGACACACAGCAAGTGAATTATGGTGAACGGTACTGCGGACCTTC 2280
Db |||||
QY 2281 AAGGTCACACCATCCAAACCGTTTGGTGGCTGCTGAGCGCAGGTATTAAACGCCCCAGT 2340
Db |||||
QY 2281 AAGGTCACACCATCCAAACCGTTTGGTGGCTGCTGAGCGCAGGTATTAAACGCCCCAGT 2340
Db |||||
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTCTCTGAAAACCTATCTGCTGACTGATGAAGTCTG 2400
Db |||||
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTCTCTGAAAACCTATCTGCTGACTGATGAAGTCTG 2400
Db |||||
QY 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGGCTGAGCGCAGGTATTAAACGCCCCAGT 2460
Db |||||
QY 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGGCTGAGCGCAGGTATTAAACGCCCCAGT 2460
Db |||||
QY 2461 TTGGCGAAAGATCCACGTTATTCGCCACCACATGGAAACCGCCAGAAAGTGAATCATG 2520
Db |||||
QY 2461 TTGGCGAAAGATCCACGTTATTCGCCACCACATGGAAACCGCCAGAAAGTGAATCATG 2520
Db |||||
QY 2521 CCGAACATCCCGCAGATGTCGCTTCTGGTATGCGGTGCGTACTGCGGCTGATCAACGCC 2580
Db |||||
QY 2521 CCGAACATCCCGCAGATGTCGCTTCTGGTATGCGGTGCGTACTGCGGCTGATCAACGCC 2580
Db |||||
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db |||||
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db |||||
QY 2641 AACACACACACATACAAATACAAACCTCGGATCGAGGAAAGATTCAGAAATTC 2700
Db |||||
QY 2641 AACACACACACATACAAATACAAACCTCGGATCGAGGAAAGATTCAGAAATTC 2700
Db |||||

RESULT 14
US-10-263-153-20
; Sequence 20, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263.153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 7370
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)....(3447)
; OTHER INFORMATION: pMBP-c2x-Toxo30del13C (52-300aa)
US-10-263-153-20
```

Query Match 81.5%; Score 2690,4; DB 18; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
QY 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGGGTATGGCATATAGCGCCGGAAGAGA 60
Db |||||
QY 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGGGTATGGCATATAGCGCCGGAAGAGA 60
Db |||||
QY 61 GTCAAATTCAGGGTGGTGAATGTGAACCAAGTAACGTTATACGATGTGCGAGATATGCGG 120
Db |||||
QY 61 GTCAAATTCAGGGTGGTGAATGTGAACCAAGTAACGTTATACGATGTGCGAGATATGCGG 120
Db |||||
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGCGCAGCCAGTTTCTCGGAAA 180
Db |||||
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGCGCAGCCAGTTTCTCGGAAA 180
Db |||||
QY 181 CGCGGAAAAGTGAAGCGCGGAGTGGCGGAGCTGAATTACATTCCAAACCGGTGGCAC 240
Db |||||
QY 181 CGCGGAAAAGTGAAGCGCGGAGTGGCGGAGCTGAATTACATTCCAAACCGGTGGCAC 240
Db |||||
QY 241 AACAACTGGCGGGCAAAACAGTTCGTTGCTGATTCGGCGTTCGACCTCCAGTCTGGCCCTGC 300
Db |||||
QY 241 AACAACTGGCGGGCAAAACAGTTCGTTGCTGATTCGGCGTTCGACCTCCAGTCTGGCCCTGC 300
Db |||||
QY 301 ACGCGCGTCCGAAATTTGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db |||||
QY 301 ACGCGCGTCCGAAATTTGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db |||||
QY 361 TGGTGGTGTGATGAGTGAACGAGCGGCTGCAAGCTGTAAAGCGCGGTGCACAATC 420
Db |||||
QY 361 TGGTGGTGTGATGAGTGAACGAGCGGCTGCAAGCTGTAAAGCGCGGTGCACAATC 420
Db |||||
QY 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTATCGCTGGATGACCAAGATGCCA 480
Db |||||
QY 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTATCGCTGGATGACCAAGATGCCA 480
Db |||||
QY 481 TTGCTGTGGAAGCTGCTGCACTAATTTCTCGCGCTTATTTCTTGATGTCTCTGACCAGA 540
Db |||||
QY 481 TTGCTGTGGAAGCTGCTGCACTAATTTCTCGCGCTTATTTCTTGATGTCTCTGACCAGA 540
Db |||||
QY 541 CACCCATCAACAGTATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db |||||
QY 541 CACCCATCAACAGTATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db |||||
QY 601 TGGTGCATTTGGTCAACGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTCTCTGCTCGG 660
Db |||||
QY 601 TGGTGCATTTGGTCAACGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTCTCTGCTCGG 660
Db |||||
QY 661 CGGCTCTGCGTCTGGCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db |||||
QY 661 CGGCTCTGCGTCTGGCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db |||||
QY 721 CGGAACGGGAAGCGGACTGGAGTGCATGTCGCGTTCGCGGTTTCAACAAACCATGCAAAATGCTGA 780
Db |||||
QY 721 CGGAACGGGAAGCGGACTGGAGTGCATGTCGCGTTCGCGGTTTCAACAAACCATGCAAAATGCTGA 780
Db |||||
QY 781 ATGAGGGCATCGTTCCCATCGGATGCTGGTTGCCAAGATGAGATGGCGTGGCGCAA 840
Db |||||
QY 781 ATGAGGGCATCGTTCCCATCGGATGCTGGTTGCCAAGATGAGATGGCGTGGCGCAA 840
Db |||||
QY 841 TGGCGCGCATTTACCGAGTCCGCGTGGTGGCGGATATCTCGGTAGTGGGATAG 900
Db |||||
QY 841 TGGCGCGCATTTACCGAGTCCGCGTGGTGGCGGATATCTCGGTAGTGGGATAG 900
Db |||||
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAACAGGATTTTC 960
Db |||||
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAACAGGATTTTC 960
Db |||||
QY 961 GCCTGCTGGGCAACACAGCGTGGAGCGCTGCTGCACTCTCTCAGGGCCAGGGCGTGA 1020
Db |||||
QY 961 GCCTGCTGGGCAACACAGCGTGGAGCGCTGCTGCACTCTCTCAGGGCCAGGGCGTGA 1020
Db |||||
QY 1021 AGGGCAATCAGTGTGTCGCTCTCACTGGTGAAGAAAACCAACCTCGCGCCCAATA 1080
Db |||||
QY 1021 AGGGCAATCAGTGTGTCGCTCTCACTGGTGAAGAAAACCAACCTCGCGCCCAATA 1080
Db |||||
QY 1081 CGCAACCGCCTCTCCCGCGGTGGCGGATTCATTAATGCGTGGCAGCAGACAGGTTT 1140
Db |||||
```

Db 1081 ||||| CGCAACCGCCTCTCCCGCGGCTTGGCGGATTCATTAATGCGAGTGGCACGACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Qy 1201 GCACAAATCTCATGTTTGACAGCTTATCATGACGTGACGGTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATGACGTGACGGTGACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGACGCATCGGAAGCTGTGATGCTGTGAGGTGCGTAAATCACTGCGATAAATTCG 1320
Db 1261 TCAGGACGCATCGGAAGCTGTGATGCTGTGAGGTGCGTAAATCACTGCGATAAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTCTCGATATGTTTTTGGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTCGATATGTTTTTGGCGCGACATCAATACGGTT 1380
Qy 1381 CTGGCAAAATCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTTGTTG 1440
Db 1381 CTGGCAAAATCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTTGTTG 1440
Qy 1441 ATTGTGAGCGGATAACAATTTTCCACAGCAACAGCCAGTCCGTTTAGTGTTTTTCAGCA 1500
Db 1441 ATTGTGAGCGGATAACAATTTTCCACAGCAACAGCCAGTCCGTTTAGTGTTTTTCAGCA 1500
Qy 1501 GCACCTTCAACCAAGGACCATAGATTAATGAAAATGGAAGGTAAGCTGTTTCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTAATGAAAATGGAAGGTAAGCTGTTTCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGTGAAGTCCGTTAAGAAATTCAGAAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAAAGGCTCTCGTGAAGTCCGTTAAGAAATTCAGAAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Qy 1681 GCGCAACTGCGGATGCGCTTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGCGGATGCGCTTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTTGGCTGAATATCAACCGGACAAAGGTTTCAGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTTGGCTGAATATCAACCGGACAAAGGTTTCAGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTTACACGCGAAGCTGATTGCTTACCGGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACACGCGAAGCTGATTGCTTACCGGATCGCTGT 1860
Qy 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTCGATAAAGAACTGAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTCGATAAAGAACTGAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGGCGGCTGATTGCTGCTGAGCGGGGTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGGCTGATTGCTGCTGAGCGGGGTTATCGGTTCAAG 2040
Qy 2041 TATGAAAACCGCAAGTACGACATTAAGACGTTGGCGGTGGATAACCGTGGCGGAAAGCG 2100
Db 2041 TATGAAAACCGCAAGTACGACATTAAGACGTTGGCGGTGGATAACCGTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACATGAATTCAGACACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACATGAATTCAGACACCGGATTAC 2160
Qy 2161 TCCATCGCAAGAGCTGCCTTTAATTAAGCGCAACAGCGATGACCATCAACGGCCCGGTG 2220

Db 2161 TCCATCGCAAGAGCTGCCTTTAATAAAGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCCAAATCATCGACACAGCAAAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAATCATCGACACAGCAAAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
Qy 2281 AAGGGTCAACCATCAAAACCGTTTGGCGGTGCTGAGCGCAGGTATTAACGGCCGCCAGT 2340
Db 2281 AAGGGTCAACCATCAAAACCGTTTGGCGGTGCTGAGCGCAGGTATTAACGGCCGCCAGT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTTCTCGAAAACATATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTTCTCGAAAACATATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAAGCGGTTAATAAAGCAAAACCGCTGGTCCGTAGCGCTGAAAGTCTTTACGAGGAAG 2460
Db 2401 GAAAGCGGTTAATAAAGCAAAACCGCTGGTCCGTAGCGCTGAAAGTCTTTACGAGGAAG 2460
Qy 2461 TTGGGGAAGATCCACGTATTTGGCGCCACATGGAACGCCAGAAAAGGTGAATCATG 2520
Db 2461 TTGGGGAAGATCCACGTATTTGGCGCCACATGGAACGCCAGAAAAGGTGAATCATG 2520
Qy 2521 CCGAAACATCCCGCAGATGTCGCTTTTCTGGTATGCGGTGCTGCTGATCAACGCC 2580
Db 2521 CCGAAACATCCCGCAGATGTCGCTTTTCTGGTATGCGGTGCTGCTGATCAACGCC 2580
Qy 2581 GCCAGCGGTCGTACAGCTGTGATGAAGCCTCGTGAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTCGTACAGCTGTGATGAAGCCTCGTGAAGACGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACACACACAAATTAACAATAACAACCAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACACACACAAATTAACAATAACAACCAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700

RESULT 15

US-10-263-153-61
; Sequence 61, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US 01
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 7370
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3447)
; OTHER INFORMATION: pMBP-c2X-Toxop30MIX1
US-10-263-153-61

Query Match 81.5%; Score 2690.4; DB 18; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CCGACCATCGAATGGTGC AAAACCTTTTCGGGTATGCGATGATAGCCCGGAAGAGA 60
Db 1 CCGACCATCGAATGGTGC AAAACCTTTTCGGGTATGCGATGATAGCCCGGAAGAGA 60
Qy 61 GTCATTCAGGGTGGTGAATGTAACCAAGTAAGCTTATACGATGTCGACAGTATCGCG 120
Db 61 GTCATTCAGGGTGGTGAATGTAACCAAGTAAGCTTATACGATGTCGACAGTATCGCG 120

QY 121 GTGTCTCTTATCAGACCGGTTTCCCGCGTGGTGAACACGCGCAGCCAGGTTTCTCGGAAA 180
Db 121 GTGTCTCTTATCAGACCGGTTTCCCGCGTGGTGAACACGCGCAGCCAGGTTTCTCGGAAA 180
QY 181 CGCGGGAAAAGTGAAGCGGGGATGGCGGAGCTGAATTAATCCCAACCGCGTGGCAC 240
Db 181 CGCGGGAAAAGTGAAGCGGGGATGGCGGAGCTGAATTAATCCCAACCGCGTGGCAC 240
QY 241 AACAACTCGCGGGCAAAACAGTCTGTGTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTCGCGGGCAAAACAGTCTGTGTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTCCGAAATTTGTCCGCGGATTAATCTCGCGCGATCAACTCGGTTGCCAGCG 360
Db 301 ACGCGCGTCCGAAATTTGTCCGCGGATTAATCTCGCGCGATCAACTCGGTTGCCAGCG 360
QY 361 TGGTGGTGTGATGCTAGAACGAGCGGCTCGAAGCTGTAAAGCGGCGGTGCAAAATC 420
Db 361 TGGTGGTGTGATGCTAGAACGAGCGGCTCGAAGCTGTAAAGCGGCGGTGCAAAATC 420
QY 421 TTCTCGCGCAACGCGTCACTGGGCTGATCAATTAATCTCGCTGGATGACCAAGGATGCCA 480
Db 421 TTCTCGCGCAACGCGTCACTGGGCTGATCAATTAATCTCGCTGGATGACCAAGGATGCCA 480
QY 481 TTGCTGTGGAGCTGCTGCACTAATGTTCCGCGTTATTTCTTGATGTCTCTGACCAGA 540
Db 481 TTGCTGTGGAGCTGCTGCACTAATGTTCCGCGTTATTTCTTGATGTCTCTGACCAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
QY 601 TGGTGCATTTGGTTCACCGCAATCGCGCTGTAGCGGGCCCAATAAGTTCTGTCTCGG 660
Db 601 TGGTGCATTTGGTTCACCGCAATCGCGCTGTAGCGGGCCCAATAAGTTCTGTCTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCGGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCGGATAG 720
QY 721 CGGACGGGAAGCGACTGGAGTGCATGTCCGTTTTCAACAAACCAATGCAATGTGA 780
Db 721 CGGACGGGAAGCGACTGGAGTGCATGTCCGTTTTCAACAAACCAATGCAATGTGA 780
QY 781 ATGAGGGCATCGTTCCCACTGGATGCTGTGGTGGCAATCAATGCAATGCTGA 840
Db 781 ATGAGGGCATCGTTCCCACTGGATGCTGTGGTGGCAATCAATGCAATGCTGA 840
QY 841 TCGCGCCATTAACGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TCGCGCCATTAACGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAACAGGATTTTC 960
QY 961 GCCTGTGGGGCAAAACGAGGTGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGTGGGGCAAAACGAGGTGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGCGCGTCTCACTGGTGAAGAAAGAAACCACTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGCGCGTCTCACTGGTGAAGAAAGAAACCACTGGCGCCCAATA 1080
QY 1081 CGGAAACCGGCTCTCCCGCGTGGCCGATTCATTAATGAGCTGGCACGACGAGTTT 1140
Db 1081 CGGAAACCGGCTCTCCCGCGTGGCCGATTCATTAATGAGCTGGCACGACGAGTTT 1140
QY 1141 CCCGACTGGAAAGCGGAGTGGCAACGCAATTAATGTGAGTTAGTCTCACTTAG 1200
Db 1141 CCCGACTGGAAAGCGGAGTGGCAACGCAATTAATGTGAGTTAGTCTCACTTAG 1200

QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGACGGTGACCAATGCTTCTGCGC 1260
Db 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGACGGTGACCAATGCTTCTGCGC 1260
QY 1261 TCAGGAGCCTATCGGAAGCTGTGGTATGGCTGTGAGGTCTGTAATCACTGCAATATTCG 1320
Db 1261 TCAGGAGCCTATCGGAAGCTGTGGTATGGCTGTGAGGTCTGTAATCACTGCAATATTCG 1320
QY 1321 TGTGCTCAAGCGCAGCTCCCGTTCTGGAATAATGTTTTTGGCGCGACATCATAAACGGTT 1380
Db 1321 TGTGCTCAAGCGCAGCTCCCGTTCTGGAATAATGTTTTTGGCGCGACATCATAAACGGTT 1380
QY 1381 CTGGCAATAATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTGTAATATGTGTGGA 1440
Db 1381 CTGGCAATAATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTGTAATATGTGTGGA 1440
QY 1441 ATTGTGAGCGGATCAAAATTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTACGA 1500
Db 1441 ATTGTGAGCGGATCAAAATTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTACGA 1500
QY 1501 GCATTTCAACAAACAGGACCATAGATTATGAAACTGGAAGGTAAACTGTGTAATCTGG 1560
Db 1501 GCATTTCAACAAACAGGACCATAGATTATGAAACTGGAAGGTAAACTGTGTAATCTGG 1560
QY 1561 ATTAAACGGGATTAAGGCTATAACGGTCTCGCTGGAAGTGGTAAGAAATTCGAGAAGAT 1620
Db 1561 ATTAAACGGGATTAAGGCTATAACGGTCTCGCTGGAAGTGGTAAGAAATTCGAGAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGTAAACTGGAAGGTAAACTGTGTAATCTGG 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGTAAACTGGAAGGTAAACTGTGTAATCTGG 1680
QY 1681 GCGGCAACTGGCGATGGCCTGACATTAATCTTTCTGGGCAACAGCCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCTGACATTAATCTTTCTGGGCAACAGCCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCGTGTGGCTGAAATCAACCGGCAAGCTGATGCTTACCCTGCTGTT 1800
Db 1741 GCTCAATCTGGCGTGTGGCTGAAATCAACCGGCAAGCTGATGCTTACCCTGCTGTT 1800
QY 1801 CGTTTACTGGGATGCGGTACCAACGGCAAGCTGATGCTTACCCTGCTGTT 1860
Db 1801 CGTTTACTGGGATGCGGTACCAACGGCAAGCTGATGCTTACCCTGCTGTT 1860
QY 1861 GAAGCGTTATCCCTCATTTATAACAAAGATCTGTGCGGAAACCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCCCTCATTTATAACAAAGATCTGTGCGGAAACCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCTGGATTAAGAACTGAAAGGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCTGGATTAAGAACTGAAAGGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACGCTACTTCACTGCGCGTGAATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACGCTACTTCACTGCGCGTGAATGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGAGCGTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGAGCGTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCCCTGCTGATTAAGAAACCAACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCCCTGCTGATTAAGAAACCAACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAGGCGGAAACAGCGATGACCATCAACGCGCGGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAGGCGGAAACAGCGATGACCATCAACGCGCGGTGG 2220
QY 2221 GCATGGTCCACATCGACACGAAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCACATCGACACGAAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTCTGGCTGTGAGCGCAGGTATTAAACGCGCCAGT 2340

[illegible]

Search completed: August 3, 2005, 04:22:00
Job time : 1317.2 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 352.4 Seconds
(without alignments)
15322.682 Million cell updates/sec

Title: US-09-765-555B-16
Perfect score: 3300
Sequence: 1 ccgacaccatcgaaatggtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10_0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2700	81.8	7475	2	US-08-971-036-1
2	2700	81.8	7475	3	US-09-096-570-1
c	3	2700	81.8	7475	Sequence 1, Appli
	4	1417.6	43.0	5926	Sequence 1, Appli
	5	1362	41.3	5201	Sequence 3, Appli
	6	1362	41.3	5201	Sequence 2, Appli
	7	1204	36.5	4557	Sequence 3, Appli
	8	1200.8	36.4	3832	Sequence 5, Appli
	9	1199.2	36.3	5248	Sequence 2, Appli
c	10	1199.2	36.3	5248	Sequence 18, Appl
c	11	1199.2	36.3	5312	Sequence 21, Appl
c	12	1199.2	36.3	5443	Sequence 35, Appl
c	13	1199.2	36.3	5502	Sequence 1, Appli
c	14	1199.2	36.3	5502	Sequence 785, App
c	15	1199.2	36.3	5502	Sequence 785, App
c	16	1199.2	36.3	5502	Sequence 785, App
c	17	1199.2	36.3	5502	Sequence 785, App
c	18	1199.2	36.3	5502	Sequence 785, App
c	19	1199.2	36.3	5616	Sequence 785, App
c	20	1199.2	36.3	5873	Sequence 62, Appli
c	21	1199.2	36.3	6353	Sequence 437A-62
c	22	1199.2	36.3	6353	Sequence 784, App
c	23	1199.2	36.3	6353	Sequence 784, App
c	24	1199.2	36.3	6353	Sequence 784, App
c	25	1199.2	36.3	6353	Sequence 784, App
c	26	1199.2	36.3	6353	Sequence 784, App
27	1199.2	36.3	6361	3	US-08-646-538-7

c	28	1199.2	36.3	6361	3	US-09-503-222-7	Sequence 7, Appli
c	29	1199.2	36.3	6727	2	US-08-125-462-2	Sequence 2, Appli
c	30	1199.2	36.3	6727	2	US-08-891-848-2	Sequence 2, Appli
c	31	1199.2	36.3	6799	2	US-08-125-462-5	Sequence 5, Appli
c	32	1199.2	36.3	6799	2	US-08-891-848-5	Sequence 5, Appli
c	33	1199.2	36.3	7659	3	US-09-128-314-4	Sequence 4, Appli
c	34	1199.2	36.3	7659	3	US-09-056-556-213	Sequence 213, App
c	35	1199.2	36.3	7676	3	US-09-072-596-208	Sequence 208, App
c	36	1199.2	36.3	7676	4	US-09-072-967-213	Sequence 213, App
c	37	1199.2	36.3	7676	4	US-09-287-849-9	Sequence 9, Appli
c	38	1199.2	36.3	8031	3	US-09-643-597-254	Sequence 254, App
c	39	1199.2	36.3	8031	4	US-09-480-884A-254	Sequence 254, App
c	40	1199.2	36.3	8031	4	US-09-542-615A-254	Sequence 254, App
c	41	1199.2	36.3	8031	4	US-09-606-421B-254	Sequence 254, App
c	42	1199.2	36.3	8031	4	US-09-476-496A-254	Sequence 254, App
c	43	1199.2	36.3	8031	4	US-09-630-940B-254	Sequence 254, App
c	44	1199.2	36.3	8157	3	US-09-128-314-3	Sequence 3, Appli
c	45	1199.2	36.3	8501	3	US-08-793-900-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-971-036-1
; Sequence 1, Application US/08971036
; Patent No. 5866684
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R
; APPLICANT: Hurst, David N
; APPLICANT: Jones, Philip S
; APPLICANT: Kay, Paul B
; APPLICANT: Raynham, Tony M
; APPLICANT: Wilson, Francis X
; TITLE OF INVENTION: Amino Acid Derivatives
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingland Street
; CITY: Nutley
; STATE: N.J
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/971,036
; FILING DATE: 14-NOV-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9623908.2
; FILING DATE: 18-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kreisler, Lewis J
; REGISTRATION NUMBER: 38522
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (973) 235-4387
; TELEFAX: (973) 235-2363
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7475 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
US-08-971-036-1

Query Match 81.8%; Score 2700; DB 2; Length 7475;

Best Local Similarity 100.0%; Pred. No. 0;				Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	CGACACCATCGAATGTCGCAAACTTTCGGGTATGCGATGATAGCCCGGAGAGA	60				
Db	1	CGACACCATCGAATGTCGCAAACTTTCGGGTATGCGATGATAGCCCGGAGAGA	60				
QY	61	GTCAATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGAGAGTATCGG	120				
Db	61	GTCAATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGAGAGTATCGG	120				
QY	121	GTGTCCTTATCAGACCGTTTCGGGTGGTGAACAGCCAGCCACGTTTCTGGAAAA	180				
Db	121	GTGTCCTTATCAGACCGTTTCGGGTGGTGAACAGCCAGCCACGTTTCTGGAAAA	180				
QY	181	CGCGGAAAAAGTGAAGCGGAGTGGCGGCTGAAATACATTCACATCCCAACCGCTGGCAC	240				
Db	181	CGCGGAAAAAGTGAAGCGGAGTGGCGGCTGAAATACATTCACATCCCAACCGCTGGCAC	240				
QY	241	AACAACTGGCGGCAACAGTCGTTGCTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGC	300				
Db	241	AACAACTGGCGGCAACAGTCGTTGCTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGC	300				
QY	301	ACGCGCGTCGCAAAATGTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGC	360				
Db	301	ACGCGCGTCGCAAAATGTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGC	360				
QY	361	TGGTGGTTCGATGTTAGAACGAGCGGCTCGAAGCTGTAAAGCGCGGTGACAAATC	420				
Db	361	TGGTGGTTCGATGTTAGAACGAGCGGCTCGAAGCTGTAAAGCGCGGTGACAAATC	420				
QY	421	TTCTCGCGAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATCACCAGGATGCCA	480				
Db	421	TTCTCGCGAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATCACCAGGATGCCA	480				
QY	481	TTGCTGTGGAAGCTCCCTGCACTATGTTCCGGGTTATTTCTTGATGTCCTGACACGA	540				
Db	481	TTGCTGTGGAAGCTCCCTGCACTATGTTCCGGGTTATTTCTTGATGTCCTGACACGA	540				
QY	541	CACCCATCAACAGATATTATTTCTCCATGAAGCGGTACGCGATCGGCGGTGAGGATC	600				
Db	541	CACCCATCAACAGATATTATTTCTCCATGAAGCGGTACGCGATCGGCGGTGAGGATC	600				
QY	601	TGGTGCATTCGGTTCACAGCAATTCGCGTGTAGCGGCGCAATTAAGTTCTGTCTCG	660				
Db	601	TGGTGCATTCGGTTCACAGCAATTCGCGTGTAGCGGCGCAATTAAGTTCTGTCTCG	660				
QY	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720				
Db	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720				
QY	721	CGGAACGGGAAGGCGACTGGAGTGCATGTCCGGTTTCAACAAACCATGCAATGCTGA	780				
Db	721	CGGAACGGGAAGGCGACTGGAGTGCATGTCCGGTTTCAACAAACCATGCAATGCTGA	780				
QY	781	ATGAGGGCATGTTCCCACTCGCATGCTGTTAGCGGCGCAATTAAGTTCTGTCTCG	840				
Db	781	ATGAGGGCATGTTCCCACTCGCATGCTGTTAGCGGCGCAATTAAGTTCTGTCTCG	840				
QY	841	TGCGCGCATTCACGAGTCCGGCTGGCGTGGTGGTGGATATCTCGGTAGTGGGATACG	900				
Db	841	TGCGCGCATTCACGAGTCCGGCTGGCGTGGTGGTGGATATCTCGGTAGTGGGATACG	900				
QY	901	ACGATACCGAGACAGCTCATGTTATTCGCGCGTTAACCAATCAACAGGATTTTC	960				
Db	901	ACGATACCGAGACAGCTCATGTTATTCGCGCGTTAACCAATCAACAGGATTTTC	960				
QY	961	GCCTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020				
Db	961	GCCTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020				
QY	1021	AGGCAATACGCTGTTCCCGCTCTCACTGTGTAAGAAAAAACAACCCCTGGCGCCCAATA	1080				
Db	1021	AGGCAATACGCTGTTCCCGCTCTCACTGTGTAAGAAAAAACAACCCCTGGCGCCCAATA	1080				
QY	1081	CGAAAAACCGCTCTCCCGCGGTTGGCCGATTCATTAATGACGTGGCAGCAGAGTTT	1140				
Db	1081	CGAAAAACCGCTCTCCCGCGGTTGGCCGATTCATTAATGACGTGGCAGCAGAGTTT	1140				
QY	1141	CCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGTCACTCATTTAG	1200				
Db	1141	CCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGTCACTCATTTAG	1200				
QY	1201	GCATAATCTCATGTTTGACAGCTTATCATCGACTGACCGGTGACCAATGCTTCTGGCG	1260				
Db	1201	GCATAATCTCATGTTTGACAGCTTATCATCGACTGACCGGTGACCAATGCTTCTGGCG	1260				
QY	1261	TCAGCAGCCATCGAAAGCTGTGATGCGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG	1320				
Db	1261	TCAGCAGCCATCGAAAGCTGTGATGCGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG	1320				
QY	1321	TGTCGCTCAAGGCGCACTCCGTTCTCGATAATGTTTTTTTGGCGCGACATCAACCGTT	1380				
Db	1321	TGTCGCTCAAGGCGCACTCCGTTCTCGATAATGTTTTTTTGGCGCGACATCAACCGTT	1380				
QY	1381	CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA	1440				
Db	1381	CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA	1440				
QY	1441	ATTGTAGCGGATTAACAATTTTCAACAGGAAACAGCCAGTCCGTTTGGTGTGTTTTCACGA	1500				
Db	1441	ATTGTAGCGGATTAACAATTTTCAACAGGAAACAGCCAGTCCGTTTGGTGTGTTTTCACGA	1500				
QY	1501	GCATTCACCAACAGGACCATAGATTATGAAACTGAAAGGTAACGTTAAATCTGG	1560				
Db	1501	GCATTCACCAACAGGACCATAGATTATGAAACTGAAAGGTAACGTTAAATCTGG	1560				
QY	1561	ATTAAACGGCGATTAAGGCTATAACGGTCTCGCTGAAAGTTCGTAAGAAATTCGAGAAAGAT	1620				
Db	1561	ATTAAACGGCGATTAAGGCTATAACGGTCTCGCTGAAAGTTCGTAAGAAATTCGAGAAAGAT	1620				
QY	1621	ACCGGAATTAAGTCAACCGTTGAGCATTCGGATTAACCTGGAAGAAATTCGCAAGGTT	1680				
Db	1621	ACCGGAATTAAGTCAACCGTTGAGCATTCGGATTAACCTGGAAGAAATTCGCAAGGTT	1680				
QY	1681	CGCGCAACTGGCGATCGCCCTGACATTAATCTCTGGGCAACAGCCGCTTGGTGCTAC	1740				
Db	1681	CGCGCAACTGGCGATCGCCCTGACATTAATCTCTGGGCAACAGCCGCTTGGTGCTAC	1740				
QY	1741	GCTCAATCTGGCGTGTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAGCTGTAT	1800				
Db	1741	GCTCAATCTGGCGTGTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAGCTGTAT	1800				
QY	1801	CGTTTTACCTGGGATCGCGTACGTTACACGGCAAGCTGATTTGCTTACCGATCGCTGTT	1860				
Db	1801	CGTTTTACCTGGGATCGCGTACGTTACACGGCAAGCTGATTTGCTTACCGATCGCTGTT	1860				
QY	1861	GAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCAAACCCGCAAAAACCTGGGAA	1920				
Db	1861	GAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCAAACCCGCAAAAACCTGGGAA	1920				
QY	1921	GAGATCCCGCGCTGATTAAGAACTGAAAGGAAAGGTAAGAGCGCGCTGTATGTTCAAC	1980				
Db	1921	GAGATCCCGCGCTGATTAAGAACTGAAAGGAAAGGTAAGAGCGCGCTGTATGTTCAAC	1980				
QY	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATTTGCTGCTGACGGGGTTATCGTTCAAG	2040				
Db	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATTTGCTGCTGACGGGGTTATCGTTCAAG	2040				
QY	2041	TATGAAACCGCAAGTACGACATTAAGAGCTGGCGTGGATTAACGCTGGCGCGAAGCG	2100				
Db	2041	TATGAAACCGCAAGTACGACATTAAGAGCTGGCGTGGATTAACGCTGGCGCGAAGCG	2100				
QY	2101	GGTCTGACCTTCTGTTGACCTGATTAATAAACAACATCAATGACGACACCGATTAC	2160				
Db	2101	GGTCTGACCTTCTGTTGACCTGATTAATAAACAACATCAATGACGACACCGATTAC	2160				

QY 2161 TCCATCGCAGAGCTGCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGCTCAACATCGACACAGCAAGTGAATTATGGTAAACGGTACTGCGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACAGCAAGTGAATTATGGTAAACGGTACTGCGACCTTC 2280
QY 2281 AAGGTCACCATCAACACCGTTGCTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTCACCATCAACACCGTTGCTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
QY 2341 CGAACAAAGAGCTGGCAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGGTCGT 2400
Db 2341 CGAACAAAGAGCTGGCAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGGTCGT 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGATTATGCCGCCACCATGGAAACCGCCAGAAAGTGAAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTATGCCGCCACCATGGAAACCGCCAGAAAGTGAAATCATG 2520
QY 2521 CGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTATGCGGTGATCAACGCC 2580
Db 2521 CGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTATGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTGTGACACTGTCGATGAAGCCCTGGAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTGTGACACTGTCGATGAAGCCCTGGAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACACAAACAAATAACAAATACCAATACCAACCTCGGATCGAGGAGGATTTCAAGATTC 2700
Db 2641 AACACAAACAAATAACAAATACCAATACCAACCTCGGATCGAGGAGGATTTCAAGATTC 2700

RESULT 2

US-09-096-570-1
; Sequence 1, Application US/09096570
; Patent No. 6018020
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R
; APPLICANT: Hurst, David N
; APPLICANT: Jones, Philip S
; APPLICANT: Kay, Paul B
; APPLICANT: Raynham, Tony M
; APPLICANT: Wilson, Francis X
; TITLE OF INVENTION: Amino Acid Derivatives
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: N.J.
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/096, 570
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/971,036
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kreibler, Lewis J
; REGISTRATION NUMBER: 38522

; REFERENCE/DOCKET NUMBER: RAN 4430/073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (973) 235-4387
; TELEFAX: (973) 235-2363
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7475 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
US-09-096-570-1

Query Match 81.8%; Score 2700; DB 3; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGGCATGTAGCGCCGGGAGAGA 60
Db 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGGCATGTAGCGCCGGGAGAGA 60
QY 61 GTC AATTCAGGGTGGTGAATGTGAAACCACTTAACGATGTCGAGAGTATGCCG 120
Db 61 GTC AATTCAGGGTGGTGAATGTGAAACCACTTAACGATGTCGAGAGTATGCCG 120
QY 121 GTGTCCTTTATCAGACCGTTTCCCGCGTGGTGAACCCAGCCAGCCACGTTTCTGCGAAAA 180
Db 121 GTGTCCTTTATCAGACCGTTTCCCGCGTGGTGAACCCAGCCAGCCACGTTTCTGCGAAAA 180
QY 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTACATTCCTCAACCGCTGGGCAC 240
Db 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTACATTCCTCAACCGCTGGGCAC 240
QY 241 AACAACTCGCGCGGCAACAGTCGTTGCTGATTTGGGGTTGCCACCTCCAGTCTGGGCCCTGC 300
Db 241 AACAACTCGCGCGGCAACAGTCGTTGCTGATTTGGGGTTGCCACCTCCAGTCTGGGCCCTGC 300
QY 301 ACGGCGCTCGCAAAATTCGCGCGGATTAATATCTCGCGCGGATCAACTGGGTGCCAGCG 360
Db 301 ACGGCGCTCGCAAAATTCGCGCGGATTAATATCTCGCGCGGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTGTGATGTTAGAAAGCGCGCTGGAAGCTGTAAAGCGCGGTGCAACAATC 420
Db 361 TGGTGGTGTGATGTTAGAAAGCGCGCTGGAAGCTGTAAAGCGCGGTGCAACAATC 420
QY 421 TTCTCGCGCAACGGCTCAGTGGGCTGATTAATTAATCTATCGCTGATGACCAAGGATGCCA 480
Db 421 TTCTCGCGCAACGGCTCAGTGGGCTGATTAATTAATCTATCGCTGATGACCAAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTTGACCCAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGGGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGGGACTGGGCGTGGAGCATC 600
QY 601 TGGTGGCATTTGGTCCACGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db 601 TGGTGGCATTTGGTCCACGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGCGGACCTGGAGTGCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGCGGACCTGGAGTGCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGGCGCAA 840
Db 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGGCGCAA 840

841 TGCGGCCATTACCGAGTCGGGCTGCGGTGGTGGGATATCTCGGTAGTGGGATACG 900
Db
841 TGCGCGCATACCGAGTCGGGCTGCGGTGGTGGGATATCTCGGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCAACCAAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCAACCAAGGATTTTC 960
Qy 961 GCTGCTGGGGCAACACGCGTGGACCGCTTCTGCACTCTCTCAGGCGCAGGCGGTGA 1020
Db 961 GCTGCTGGGGCAACACGCGTGGACCGCTTCTGCACTCTCTCAGGCGCAGGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGCTGCGCACACAGGTTT 1140
Db 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGCTGCGCACACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTCAGCGCAACGCAATTAATGTGAGTTAGCTCACTATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTCAGCGCAACGCAATTAATGTGAGTTAGCTCACTATTAG 1200
Qy 1201 GCACAATTCATGTTTGACAGCTTATCATGACTGCACGGTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTCATGTTTGACAGCTTATCATGACTGCACGGTGACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGCGCATTCGGAAGCTGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
Db 1261 TCAGGCGCATTCGGAAGCTGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTTCGATTAATGTTTTCGCGGACATCATACCGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTTCGATTAATGTTTTCGCGGACATCATACCGGTT 1380
Qy 1381 CTGGCAATATCTGAATGAGCTGTTGACATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAATATCTGAATGAGCTGTTGACATTAATCATCGGCTCGTATATGTTGGA 1440
Qy 1441 ATTGTAGCGGATTAACATTTTACAGGAAACAGCGAGCTCGGTTAGGTTTTCACGA 1500
Db 1441 ATTGTAGCGGATTAACATTTTACAGGAAACAGCGAGCTCGGTTAGGTTTTCACGA 1500
Qy 1501 GCATTTACCAACAGGACCATAGATTATGAAACTGAAGAGTAACTGGTAACTCGG 1560
Db 1501 GCATTTACCAACAGGACCATAGATTATGAAACTGAAGAGTAACTGGTAACTCGG 1560
Qy 1561 ATTAACCGCGATAAAGGCTAAACGCTCGCTGAGTTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACCGCGATAAAGGCTAAACGCTCGCTGAGTTCGGTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACGCTTGAGCTCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGCTTGAGCTCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Qy 1681 GCGCACTGCGGATGCGGCTGACATTAATCTTGGGCAACAGCGCTTGGTGGCTAC 1740
Db 1681 GCGCACTGCGGATGCGGCTGACATTAATCTTGGGCAACAGCGCTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTTGGCTGAAATCACCCTGGACAAAGCGTTCCAGGCAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTTGGCTGAAATCACCCTGGACAAAGCGTTCCAGGCAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGATGCGGCTACGTTACACGCGCAAGCTGATTCACCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGGCTACGTTACACGCGCAAGCTGATTCACCGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATACAAAGATCTGCTGCGCAACCGCGCAAAACCTCGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATACAAAGATCTGCTGCGCAACCGCGCAAAACCTCGGAA 1920

Qy 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAAACGCAAGTACGACATTAAGAGAGTGGGCGTGGATTAACGCTGGCGGCAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGAGAGTGGGCGTGGATTAACGCTGGCGGCAAGCG 2100
Qy 2101 GGTCTGACCTTCCCTGGTGGACCTGATTAAGAAACAAACACATGATGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCCCTGGTGGACCTGATTAAGAAACAAACACATGATGAGACACCGATTAC 2160
Qy 2161 TCCATCCGAGAAGCTGCTTTTAAAGAGCGAAACAGCGATGACCATCAACGCGCCGCTGG 2220
Db 2161 TCCATCCGAGAAGCTGCTTTTAAAGAGCGAAACAGCGATGACCATCAACGCGCCGCTGG 2220
Qy 2221 GCATGCTCAACATCGACACAGCAAGTGAATTTATGTTGTAACGGTACTGCGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACAGCAAGTGAATTTATGTTGTAACGGTACTGCGACCTTC 2280
Qy 2281 AAGGTCACCAACATCCCAACCGTTCGTTGGCGTCTGAGCGCAGGTATTAAACGCGCCAGT 2340
Db 2281 AAGGTCACCAACATCCCAACCGTTCGTTGGCGTCTGAGCGCAGGTATTAAACGCGCCAGT 2340
Qy 2341 CCAGCAACAGAGCTGGCAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTTCTG 2400
Db 2341 CCAGCAACAGAGCTGGCAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTTCTG 2400
Qy 2401 GAAGCGTTTAAGAGACAAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Db 2401 GAAGCGTTTAAGAGACAAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Qy 2461 TTGGCAAGAGTCCAGTATTGGCGCCACATCGGAAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCAAGAGTCCAGTATTGGCGCCACATCGGAAACCGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCAGCAACCGCAGATGTCGGCTTTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCAGCAACCGCAGATGTCGGCTTTCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Qy 2581 GCGAGCGCTGCTCAGACTGTCGATGAGCGCTGGAAGACCGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCGAGCGCTGCTCAGACTGTCGATGAGCGCTGGAAGACCGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACCAACAAACAAATAACAACTAACCAACCTCGGATCGAGGAGGATTTTCAGAAATTC 2700
Db 2641 AACCAACAAACAAATAACAACTAACCAACCTCGGATCGAGGAGGATTTTCAGAAATTC 2700

RESULT 3

US-09-265-617B-1
; Sequence 1, Application US/09265617B
; Patent No. 6372883
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R.
; APPLICANT: Hurst, David N.
; APPLICANT: Jones, Philip S.
; APPLICANT: Kay, Paul B.
; APPLICANT: Raynham, Tony M.
; APPLICANT: Wilson, Francis X.
; TITLE OF INVENTION: Antiviral Medicaments
; FILE REFERENCE: 20052 antiviral medicaments
; CURRENT APPLICATION NUMBER: US/09/265,617B
; CURRENT FILING DATE: 1999-03-10
; PRIOR APPLICATION NUMBER: GB9806815.8
; PRIOR FILING DATE: 1998-03-30
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1

; LENGTH: 7475									
; TYPE: DNA									
; ORGANISM: Artificial Sequence									
; FEATURE:									
; OTHER INFORMATION: Description of Artificial Sequence: Vector and									
; OTHER INFORMATION: Gene Fragments									
US-09-265-617B-1									
Query Match 81.8%; Score 2700; DB 3; Length 7475;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	CGGACACCATCGAATGGTGGCAAAACCTTTCCGGGTATGCGCATGATAGCGCCGGAGAGA	60						
Db	1	CGGACACCATCGAATGGTGGCAAAACCTTTCCGGGTATGCGCATGATAGCGCCGGAGAGA	60						
Qy	61	GTCAATTACAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGCAGAGTATGCCG	120						
Db	61	GTCAATTACAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGCAGAGTATGCCG	120						
Qy	121	GTGTCCTTTATCAGACCGGTTTCCCGCGTGGTGAACACGAGCCAGCCACGTTTCTGCGAAAA	180						
Db	121	GTGTCCTTTATCAGACCGGTTTCCCGCGTGGTGAACACGAGCCAGCCACGTTTCTGCGAAAA	180						
Qy	181	CGCGGAAAAAGTGAAGCGGCGATGCGGAGCTGAATTTACATTCGCCAACCGCGTGGCAC	240						
Db	181	CGCGGAAAAAGTGAAGCGGCGATGCGGAGCTGAATTTACATTCGCCAACCGCGTGGCAC	240						
Qy	241	AACAACTGCGGGCAAAACAGTCTGTTGCTGATTGGGGTTGCCACCTCCAGTCTGGGCCCTGC	300						
Db	241	AACAACTGCGGGCAAAACAGTCTGTTGCTGATTGGGGTTGCCACCTCCAGTCTGGGCCCTGC	300						
Qy	301	ACGCGCCGTTCGCAAAATGTCGCGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG	360						
Db	301	ACGCGCCGTTCGCAAAATGTCGCGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG	360						
Qy	361	TGGTGGTTCGATGTAGAAACGAAAGCGGCTGAAAGCGCTGTAAAGCGGCGTGCACAATC	420						
Db	361	TGGTGGTTCGATGTAGAAACGAAAGCGGCTGAAAGCGCTGTAAAGCGGCGTGCACAATC	420						
Qy	421	TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATCTCGCGCCGATCAACTGGGTGCCAG	480						
Db	421	TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATCTCGCGCCGATCAACTGGGTGCCAG	480						
Qy	481	TTGCTGTGGAAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTGATGTTCTCTGACCAGA	540						
Db	481	TTGCTGTGGAAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTGATGTTCTCTGACCAGA	540						
Qy	541	CACCCATCAACAGTATTTATTTCTCCCATGAAGACGTTACGGCACTGGGCGTGGAGCATC	600						
Db	541	CACCCATCAACAGTATTTATTTCTCCCATGAAGACGTTACGGCACTGGGCGTGGAGCATC	600						
Qy	601	TGGTGCAGTTGGTCCACAGCAAACTCGGCTGTAGCGGGCCCATTAAGTCTGTCCTGG	660						
Db	601	TGGTGCAGTTGGTCCACAGCAAACTCGGCTGTAGCGGGCCCATTAAGTCTGTCCTGG	660						
Qy	661	CGCGTCTGGCTGTGGCTGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
Db	661	CGCGTCTGGCTGTGGCTGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
Qy	721	CGGAAACGGGAACGCGACTGGAGTGCATGTCGCGTTTTCAACAAACCATGCAAAATGCTGA	780						
Db	721	CGGAAACGGGAACGCGACTGGAGTGCATGTCGCGTTTTCAACAAACCATGCAAAATGCTGA	780						
Qy	781	ATGAGGGCATCGTCCCACTGCGATGCTGTTGCCAACGATCAGATGGCGTGGCGGCA	840						
Db	781	ATGAGGGCATCGTCCCACTGCGATGCTGTTGCCAACGATCAGATGGCGTGGCGGCA	840						
Qy	841	TGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG	900						
Db	841	TGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG	900						
Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCCATCAAAAGGATTTTC	960						

Db	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCCATCAAAAGGATTTTC	960						
Qy	961	GCCTGCTGGGGCAAAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGGCGTGA	1020						
Db	961	GCCTGCTGGGGCAAAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGGCGTGA	1020						
Qy	1021	AGGGCAATCAGCTGTTGCCCGTCTCATCTGTTGAAAGAAAAACCAACCTTGGGGCCCAATA	1080						
Db	1021	AGGGCAATCAGCTGTTGCCCGTCTCATCTGTTGAAAGAAAAACCAACCTTGGGGCCCAATA	1080						
Qy	1081	CGAAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATCAGCTGCGACGACAGGTTT	1140						
Db	1081	CGAAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATCAGCTGCGACGACAGGTTT	1140						
Qy	1141	CCGACTTGAAGAGGGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200						
Db	1141	CCGACTTGAAGAGGGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200						
Qy	1201	GCACAAATTCATGTTTGACAGCTTATCATCGCTGACCGTGACCAATGCTTCTGCGG	1260						
Db	1201	GCACAAATTCATGTTTGACAGCTTATCATCGCTGACCGTGACCAATGCTTCTGCGG	1260						
Qy	1261	TCAGGACGCAATCGGAAGCTGTGGTATGCTGTGAGGTGTAATCACTGCATAAATTCG	1320						
Db	1261	TCAGGACGCAATCGGAAGCTGTGGTATGCTGTGAGGTGTAATCACTGCATAAATTCG	1320						
Qy	1321	TGTGCTCAAGCGGCACCTCCCGTCTTGATTAATGTTTTTGGCGCGACATCAATACGGTT	1380						
Db	1321	TGTGCTCAAGCGGCACCTCCCGTCTTGATTAATGTTTTTGGCGCGACATCAATACGGTT	1380						
Qy	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTAAATGTGTGA	1440						
Db	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTAAATGTGTGA	1440						
Qy	1441	ATTGTGACGGATTAACAAATTTACACAGGAAACAGCAGTCCGTTTAGTGTTCACGA	1500						
Db	1441	ATTGTGACGGATTAACAAATTTACACAGGAAACAGCAGTCCGTTTAGTGTTCACGA	1500						
Qy	1501	GCATTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGGTAATACTGGTAATCTGG	1560						
Db	1501	GCATTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGGTAATACTGGTAATCTGG	1560						
Qy	1561	ATTAACCGCGATAAAGCTATAACCGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT	1620						
Db	1561	ATTAACCGCGATAAAGCTATAACCGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT	1620						
Qy	1621	ACCGGAATTAAGTACACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT	1680						
Db	1621	ACCGGAATTAAGTACACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT	1680						
Qy	1681	GCGCAACTGGCGATGGCCCTGACATTAATCTTTGGGCAACAGCCGCTTTTGGTGGCTAC	1740						
Db	1681	GCGCAACTGGCGATGGCCCTGACATTAATCTTTGGGCAACAGCCGCTTTTGGTGGCTAC	1740						
Qy	1741	GCTCAATCTGCGCTTGTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGAT	1800						
Db	1741	GCTCAATCTGCGCTTGTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGAT	1800						
Qy	1801	CGGTTTACCTGGGATGCGGTAGCTTACAAACGCAAGCTGATTTACCGATCGCTGTT	1860						
Db	1801	CGGTTTACCTGGGATGCGGTAGCTTACAAACGCAAGCTGATTTACCGATCGCTGTT	1860						
Qy	1861	GAAGGGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA	1920						
Db	1861	GAAGGGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA	1920						
Qy	1921	GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980						
Db	1921	GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980						
Qy	1981	CTGGAAGAACGTAATTCACCTGGCGCTGATGCTGCTGACGCGGGGTTATCGGTTCAAG	2040						

Db 1981 CTGCAAGAACCGTACTTCCCTACCTGGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAACGGCAAGTACGACATTAAGACGTTGGCGTGGATAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACGGCAAGTACGACATTAAGACGTTGGCGTGGATAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGCTGACCTGATTAATAAACAACAACATGATGACGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGACCTGATTAATAAACAACAACATGATGACGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACGCGCGGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACGCGCGGTGG 2220
Qy 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTAATGTTAAGCGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTAATGTTAAGCGTACTGCGGACCTTC 2280
Qy 2281 AAGGTCAACCATCAAAACCGTTCTGCGGTGCTGAGCGAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGTCAACCATCAAAACCGTTCTGCGGTGCTGAGCGAGGTATTAACGCGCCAGT 2340
Qy 2341 CGAAACAAAGAGCTGCGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CGAAACAAAGAGCTGCGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGTTTAAATAAGACAAACCGCTGGGTGCGGTGCTGAGCGAGGTATTAACGCGCCAGT 2460
Db 2401 GAAGCGTTTAAATAAGACAAACCGCTGGGTGCGGTGCTGAGCGAGGTATTAACGCGCCAGT 2460
Qy 2461 TTGGCGAAGATCCAGCTATTGCGCCACCATGGAAGAACGCGCCAGAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCAGCTATTGCGCCACCATGGAAGAACGCGCCAGAGGTGAATCATG 2520
Qy 2521 CCGAACATCCGCGAGATGCTCGCTTTCTGTTGATGCGGTGCTGAGCGAGGTATTAACGCGCCAGT 2580
Db 2521 CCGAACATCCGCGAGATGCTCGCTTTCTGTTGATGCGGTGCTGAGCGAGGTATTAACGCGCCAGT 2580
Qy 2581 GCCAGCGTCTGAGCTGATGATGAAGCCCTGAAAGACGCGCAGAGCTAATTCGAGTCTG 2640
Db 2581 GCCAGCGTCTGAGCTGATGATGAAGCCCTGAAAGACGCGCAGAGCTAATTCGAGTCTG 2640
Qy 2641 AACACACACAAATACAAATAACAAACCTCGGATCGAGGAGGATTTTCAAGATTC 2700
Db 2641 AACACACACAAATACAAATAACAAACCTCGGATCGAGGAGGATTTTCAAGATTC 2700

RESULT 4

US-09-027-169-3/c
Sequence 3, Application US/09027169
Patent No. 6420524
GENERAL INFORMATION:
APPLICANT: CRAIG, NANCY L
TITLE OF INVENTION: GAIN OF FUNCTION MUTATIONS IN
TITLE OF INVENTION: ATP-DEPENDENT TRANSPORT PROTEINS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Anne Brown (Alston & Bird, LLP)
STREET: 3605 Glenwood Ave.
CITY: Raleigh
STATE: NC
COUNTRY: USA
ZIP: 27608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/027,169
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: Brown, Anne
REGISTRATION NUMBER: 36,463
REFERENCE/DOCKET NUMBER: 5789-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919 420 2205
TELEFAX: 919 881 3175
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 5926 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: circular
MOLECULE TYPE: other nucleic acid
DESCRIPTION: /desc = "pEM delta R.adj to 1"
US-09-027-169-3

Query Match 43.0%; Score 1417.6; DB 3; Length 5926;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1455; Conservative 0; Mismatches 9; Indels 12; Gaps 2;

Qy 3 GACACCATCGAATGCTGCAAAACCTTTTCGCGGTATGCGCATGATAGCGCCCGGAAGAGT 62
Db 1564 GACACCATCGAATGCTGCAAAACCTTTTCGCGGTATGCGCATGATAGCGCCCGGAAGAGT 1505
Qy 63 CAATTCAAGGTGCTGTAATGTGAACAGTAACGTTATACGATGTCGAGAGTATGCGGT 122
Db 1504 CAATTCAAGGTGCTGTAATGTGAACAGTAACGTTATACGATGTCGAGAGTATGCGGT 1445
Qy 123 GTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGCGCAGCCACGTTTCTGCGAAACG 182
Db 1444 GTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGCGCAGCCACGTTTCTGCGAAACG 1385
Qy 183 CGGGAAGAGTGAACGCGCGATGCGGAGCTGAATTAATTCCTCCAAACGCGTGGCAAA 242
Db 1384 CGGGAAGAGTGAACGCGCGATGCGGAGCTGAATTAATTCCTCCAAACGCGTGGCAAA 1325
Qy 243 CAACTCGCGCGCAACAGTCTGTTGTTGTTGGCGTTGCCACCTCCAGTCTGGCCCTGCAC 302
Db 1324 CAACTCGCGCGCAACAGTCTGTTGTTGTTGGCGTTGCCACCTCCAGTCTGGCCCTGCAC 1265
Qy 303 GCGCCGTGCGAAATTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCGACGCTG 362
Db 1264 GCGCCGTGCGAAATTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCGACGCTG 1205
Qy 363 GTGCTGTCGATGTAACGAGCGCGTGAAGCTGTAAAGCGCGGTGACATCTT 422
Db 1204 GTGCTGTCGATGTAACGAGCGCGTGAAGCTGTAAAGCGCGGTGACATCTT 1145
Qy 423 CTCGCGCAACGCGTCAAGTGGGCTGATCAATTAATCTATCCGCTGGATGACCGAGTGCATT 482
Db 1144 CTCGCGCAACGCGTCAAGTGGGCTGATCAATTAATCTATCCGCTGGATGACCGAGTGCATT 1085
Qy 483 GCTGTGGAAGCTGCGTCAATTAATGTTCCGGGTTATTTCTTGATGTCCTGACGAGACA 542
Db 1084 GCTGTGGAAGCTGCGTCAATTAATGTTCCGGGTTATTTCTTGATGTCCTGACGAGACA 1025
Qy 543 CCCATCAACAGTATTAATTTCTCCCATGAAGAGCGTACGCGACTGGCGGTGAGGATCTG 602
Db 1024 CCCATCAACAGTATTAATTTCTCCCATGAAGAGCGTACGCGACTGGCGGTGAGGATCTG 965
Qy 603 GTCGCAATGGGTCAACGCAAAATCGCGTGTGTAGCGGGCCCAATTAAGTTCTGCTCGGG 662
Db 964 GTCGCAATGGGTCAACGCAAAATCGCGTGTGTAGCGGGCCCAATTAAGTTCTGCTCGGG 905
Qy 663 CGTCTGCGTCTGCGTGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
Db 904 CGTCTGCGTCTGCGTGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATAGCG 845
Qy 723 GAACGGGAAGCGACTGGAGTGCCATGTCGCGGTTTTTCAACAAACCATGCAATGCTGAAT 782
Db 844 GAACGGGAAGCGACTGGAGTGCCATGTCGCGGTTTTTCAACAAACCATGCAATGCTGAAT 785
Qy 783 GAGGCGATCGTTCCCACTCGCATGCTGTTGCGCAACGATCAGATGGCGTGGCGCAATG 842

```
|||||
Db 784 GAGGGCATGTTCCCACTGCGATGCTGGTTGGCAACGATCAGATGGCGTGGCGCAATG 725
Qy 843 CGCGCATTTACCGAGTCCGGGCTGGCGGCTGGTGGCGGATATCTCGGTAGTGGATACGAC 902
Db 724 CGCGCATTTACCGAGTCCGGGCTGGCGGCTGGTGGCGGATATCTCGGTAGTGGATACGAC 665
Qy 903 GATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAACAGAGATTTTCG 962
Db 664 GATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAACAGAGATTTTCG 605
Qy 963 CTGCTGGGGCAACACAGCTGACCGCTTGTCTGCAACTCTCTCAGGGCCAGCGGTGAAG 1022
Db 604 CTGCTGGGGCAACACAGCTGACCGCTTGTCTGCAACTCTCTCAGGGCCAGCGGTGAAG 545
Qy 1023 GCGAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAGAAACCAACCCCTGGCGCCCAATACG 1082
Db 544 GCGAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAGAAACCAACCCCTGGCGCCCAATACG 485
Qy 1083 CAAACCGCTCTCCCGCGCGTGGCGGATTAATTAATGACAGCTGGCAGCAGGTTTCC 1142
Db 484 CAAACCGCTCTCCCGCGCGTGGCGGATTAATTAATGACAGCTGGCAGCAGGTTTCC 425
Qy 1143 CGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTCATTAGGC 1202
Db 424 CGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGC-----GC 376
Qy 1203 ACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAAGCTGACCAATGCTTCTGGCGTC 1262
Db 375 GAATTTGATCTGTTTGACAGCTTATCATCGACTGCAAGCTGACCAATGCTTCTGGCGTC 316
Qy 1263 AGGCAAGCATCGGAAGCTGTGATGCTGCTGAGTGTGCGAGTGTGTAATCACTGCATAATTCGTG 1322
Db 315 AGGCAAGCATCGGAAGCTGTGATGCTGCTGAGTGTGCGAGTGTGTAATCACTGCATAATTCGTG 256
Qy 1323 TCGCTCAAGGCGCATCCCGTCTCGATAATGTTTTTGGCGCGACATCAACCGTTCT 1382
Db 255 TCGCTCAAGGCGCATCCCGTCTCGATAATGTTTTTGGCGCGACATCAACCGTTCT 196
Qy 1383 GCGAATATTTCTGAATGAGCTGTTGACAAATTAATCAT- CGGCTCGTATAATGTTGGAA 1441
Db 195 GCGAATATTTCTGAATGAGCTGTTGACAAATTAATCATCCGCTCGTATAATGTTGGAA 136
Qy 1442 TTGTGAGCGGATAACAAATTTTCACAGGAAACAGCC 1477
Db 135 TTGTGAGCGGATAACAAATTTTCACAGGAAACAGAC 100
```

RESULT 5

```
US-09-640-882-2
; Sequence 2, Application US/09640882
; Patent No. 6720142
; GENERAL INFORMATION:
; APPLICANT: Hall, Barry G.
; TITLE OF INVENTION: METHOD OF DETERMINING EVOLUTIONARY POTENTIAL OF MUTANT
; TITLE OF INVENTION: RESISTANCE GENES AND USE THEREOF TO SCREEN FOR DRUG
; TITLE OF INVENTION: EFFICACY
; FILE REFERENCE: 176/60851
; CURRENT APPLICATION NUMBER: US/09/640,882
; CURRENT FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/149,813
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 5201
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid
; OTHER INFORMATION: PACSE2
US-09-640-882-2
```

```
Query Match 41.3%; Score 1362; DB 4; Length 5201;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1380; Conservative 0; Mismatches 10; Indels 2; Gaps 1;

Qy 3 GACACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATATAGTCGCGCGGAAAGAGT 62
Db 866 GACACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATATAGTCGCGCGGAAAGAGT 925
Qy 63 CAATTCAGGGTGGTGAATGTCAAAACCACTAAACGTTATACGATGTGCGAGAGTATCCCGT 122
Db 926 CAATTCAGGGTGGTGAATGTGAAACCACTAAACGTTATACGATGTGCGAGAGTATCCCGT 985
Qy 123 GTCTCTTATCAGACGTTTCCCGGTGTGAACCAAGCCAGCCAGCTTCTTCGGAACAG 182
Db 986 GTCTCTTATCAGACGTTTCCCGGTGTGAACCAAGCCAGCCAGCTTCTTCGGAACAG 1045
Qy 183 CGGGAAGAAAGTGGAAAGCGCGATGCGGAGCTGAATTAATATATTTCCCAACCGGTGCACAA 242
Db 1046 CGGGAAGAAAGTGGAAAGCGCGATGCGGAGCTGAATTAATATATTTCCCAACCGGTGCACAA 1105
Qy 243 CAATTCAGGGTGGTGAATGTCAAAACCACTAAACGTTATACGATGTGCGAGTCTGGCCCTGCAC 302
Db 1106 CAATTCAGGGTGGTGAATGTCAAAACCACTAAACGTTATACGATGTGCGAGTCTGGCCCTGCAC 1165
Qy 303 GCGCGTGCAGAAATGTGCGCGGATTAATATCTCGCGCGATCAACTGGGTGCCAGCGT 362
Db 1166 GCGCGTGCAGAAATGTGCGCGGATTAATATCTCGCGCGATCAACTGGGTGCCAGCGT 1225
Qy 363 GTGGTGTGATGTTAGAAAGCGGCGTCAAGGCTGTAAAGCGGCGGTGCACAATCTT 422
Db 1226 GTGGTGTGATGTTAGAAAGCGGCGTCAAGGCTGTAAAGCGGCGGTGCACAATCTT 1285
Qy 423 CTCGCGCAACCGTCACTGGGCTGATCAATTAATCTACGCTGGATGACAGGATGCCATT 482
Db 1286 CTCGCGCAACCGTCACTGGGCTGATCAATTAATCTACGCTGGATGACAGGATGCCATT 1345
Qy 483 GCTGTGGAAGTGCCTGCATTAATGTTCCGGGTATTTCTTGATGTCTCTGACAGACA 542
Db 1346 GCTGTGGAAGTGCCTGCATTAATGTTCCGGGTATTTCTTGATGTCTCTGACAGACA 1405
Qy 543 CCCATCAACAGTATTATTTCTCCCATGAAGCGGTACGCACTGGCGGTGGAGCATCTG 602
Db 1406 CCCATCAACAGTATTATTTCTCCCATGAAGCGGTACGCACTGGCGGTGGAGCATCTG 1465
Qy 603 CTCGCATTGGGTCAACAGCAAAATCGCGTGTGTAGCGGCGCCATTAAGTTCTGTCTCGCG 662
Db 1466 CTCGCATTGGGTCAACAGCAAAATCGCGTGTGTAGCGGCGCCATTAAGTTCTGTCTCGCG 1525
Qy 663 CGTCTGCGTCTGGCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
Db 1526 CGTCTGCGTCTGGCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 1585
Qy 723 GAAAGGGAAGCGCTGGAGTGCATGTCGCGTTTCAACAAACCATGCAAAATGCTGAAT 782
Db 1586 GAAAGGGAAGCGCTGGAGTGCATGTCGCGTTTCAACAAACCATGCAAAATGCTGAAT 1645
Qy 783 GAGGCGATCGTTCCCATCTGCGATGCTGTGTGCAACGATCAGATGCGGTGGCGCAATG 842
Db 1646 GAGGCGATCGTTCCCATCTGCGATGCTGTGTGCAACGATCAGATGCGGTGGCGCAATG 1705
Qy 843 CGCGCATTTACGAGTCCGGCTGGCGGTGGTGGGATATCTCGGTAGTGGGATACGAC 902
Db 1706 CGCGCATTTACGAGTCCGGCTGGCGGTGGTGGGATATCTCGGTAGTGGGATACGAC 1765
Qy 903 GATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAACAGGATTTTCG 962
Db 1766 GATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAACAGGATTTTCG 1825
Qy 963 CTGCTGGGGCAACACAGCGGTGGAACCGCTTGTGCAACTCTCTCAGGGCGCAGCGGTGAAG 1022
Db 1826 CTGCTGGGGCAACACAGCGGTGGAACCGCTTGTGCAACTCTCTCAGGGCGCAGCGGTGAAG 1885
Qy 1023 GCGAATCAGCTGTTGGCGCGTCTCACTGGTGAAGAAAGAAACCAACCCCTGGCGCCCAATACG 1082
```



```
Db 1886 GGCATCAGCTGTGCGCGTCTCACTGGTGAAGAAAAACCACTGGCGCCCAATAGC 1945
Qy 1083 CAAACCGCCTCTCCCGCGCGTGGCCGATTCAATTAATGCAGCTGGCAGCAGGTTTCC 1142
Db 1946 CAAACCGCCTCTCCCGCGCGTGGCCGATTCAATTAATGCAGCTGGCAGCAGGTTTCC 2005
Qy 1143 CGACTGGAAGCGGCGAGTGAGCGCAACGAATTAATGTAGTTAGCTCACTCATTAGGC 1202
Db 2006 CGACTGGAAGCGGCGAGTGAGCGCAACGAATTAATGTAGTTAG--CGCGAATTGATC 2063
Qy 1203 ACAATCTCATGTTTGACAGCTTATCATGCTGACAGTGCACGCTGCACCAATGCTTCTGGCGTC 1262
Db 2064 TGAATCTCATGTTTGACAGCTTATCATGCTGACAGTGCACGCTGCACCAATGCTTCTGGCGTC 2123
Qy 1263 AGCGAGCCATCGGAAGCTGTGGTATGGCTGTGCAGTCTGTAATCACTGCATTAATTCGTG 1322
Db 2124 AGCGAGCCATCGGAAGCTGTGGTATGGCTGTGCAGTCTGTAATCACTGCATTAATTCGTG 2183
Qy 1323 TCGCTCAAGGCGCACTCCGTTCTGGATAAATGTTTTTGGCGCGACATCAATACGGTTCT 1382
Db 2184 TCGCTCAAGGCGCACTCCGTTCTGGATAAATGTTTTTGGCGCGACATCAATACGGTTCT 2243
Qy 1383 GGCATAATTTCT 1394
Db 2244 GGCATAATTTCT 2255

RESULT 6
US-09-640-882-3
; Sequence 3, Application US/09640882
; Patent No. 6720142
; GENERAL INFORMATION:
; APPLICANT: Hall, Barry G.
; TITLE OF INVENTION: METHOD OF DETERMINING EVOLUTIONARY POTENTIAL OF MUTANT
; TITLE OF INVENTION: RESISTANCE GENES AND USE THEREOF TO SCREEN FOR DRUG
; TITLE OF INVENTION: EFFICACY
; FILE REFERENCE: 176/60851
; CURRENT APPLICATION NUMBER: US/09/640,882
; CURRENT FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/149,813
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 5201
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid
US-09-640-882-3

Query Match 41.3%; Score 1362; DB 4; Length 5201;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1380; Conservative 0; Mismatches 10; Indels 2; Gaps 1;

Qy 3 GACACCATCGAATGGTGCNAACCTTTTCGGGTATGCGATGATAGCGCCCGAAGAGT 62
Db 866 GACACCATCGAATGGCGCAAAACCTTTTCGGGTATGCGATGATAGCGCCCGAAGAGT 925
Qy 63 CAATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGAGAGTATCGCGT 122
Db 926 CAATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGAGAGTATCGCGT 985
Qy 123 GTCTTTATCAGACCGTTTCCCGGTGGTGAACAGCCGCAACGTTTCTCGGAAAAACG 182
Db 986 GTCTTTATCAGACCGTTTCCCGGTGGTGAACAGCCGCAACGTTTCTCGGAAAAACG 1045
Qy 183 CGGGAAGAGTGGAAAGCGGAGTGGCGAGCTGTAATTAATCCACCGCTGGCAGAA 242
Db 1046 CGGGAAGAGTGGAAAGCGGAGTGGCGAGCTGTAATTAATCCACCGCTGGCAGAA 1105
```

```
Qy 243 CAACCTGGGGCAACAGTCTGTTGCTGATTTGGCGTTGGCACCTCCAGTCTGGCCCTCGAC 302
Db 1106 CAACCTGGGGCAACAGTCTGTTGCTGATTTGGCGTTGGCACCTCCAGTCTGGCCCTCGAC 1165
Qy 303 GCGCGCTGCGAAATTTGTTCGGCGCGATTAAATCTCGCCCGGATCAATCGGGTGGCAGCGTG 362
Db 1166 GCGCGCTGCGAAATTTGTTCGGCGCGATTAAATCTCGCCCGGATCAATCGGGTGGCAGCGTG 1225
Qy 363 GTGGTGTGATGTAGTAAAGCGGCTGGAAGCGCTGTAAGCGGCGGTGCACAATCTT 422
Db 1226 GTGGTGTGATGTAGTAAAGCGGCTGGAAGCGGCTGTAAGCGGCGGTGCACAATCTT 1285
Qy 423 CTGCGCAACGCGTCACTGGGCTGATCAATTAATCTATCCGCTGGATGACAGGATGCCATT 482
Db 1286 CTGCGCAACGCGTCACTGGGCTGATCAATTAATCTATCCGCTGGATGACAGGATGCCATT 1345
Qy 483 GTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTAATTTCTTGATGTCTCTGACCAGACA 542
Db 1346 GTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTAATTTCTTGATGTCTCTGACCAGACA 1405
Qy 543 CCCATCAACAGTATTATTTCTCCATGAAGACGCTACGCGACTGGGCGTGGAGCATCTG 602
Db 1406 CCCATCAACAGTATTATTTCTCCATGAAGACGCTACGCGACTGGGCGTGGAGCATCTG 1465
Qy 603 GTGCAATTTGGGTCAACAGCAATCGCGCTGTTAGCGGCGCCATTAAAGTTCTGTCGGCG 662
Db 1466 GTGCAATTTGGGTCAACAGCAATCGCGCTGTTAGCGGCGCCATTAAAGTTCTGTCGGCG 1525
Qy 663 CGTCTGCGTCTGCGTGGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
Db 1526 CGTCTGCGTCTGCGTGGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 1585
Qy 723 GAACGGGAAGCGCAGTGGAGTGCATGTCGGTTTCAACAAACCATGCAATTCGTGAAT 782
Db 1586 GAACGGGAAGCGCAGTGGAGTGCATGTCGGTTTCAACAAACCATGCAATTCGTGAAT 1645
Qy 783 GAGGCGATCGTCCCACTGCGATGCTGGTTGCAACGATCAGATGCGCTGGGCGCAATG 842
Db 1646 GAGGCGATCGTCCCACTGCGATGCTGGTTGCAACGATCAGATGCGCTGGGCGCAATG 1705
Qy 843 CGGCGCATTAACGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACGAC 902
Db 1706 CGGCGCATTAACGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACGAC 1765
Qy 903 GATACCGAAGACAGCTCATGTTTATATCCCGCGTTTAAACCAACCATCAACAGGATTTCCG 962
Db 1766 GATACCGAAGACAGCTCATGTTTATATCCCGCGTTTAAACCAACCATCAACAGGATTTCCG 1825
Qy 963 CTGCTGGGCAACACAGCGTGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGAAG 1022
Db 1826 CTGCTGGGCAACACAGCGTGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGAAG 1885
Qy 1023 GGAATACAGCTGTGCGCGCTCTCACTGGTGAAGAAAAACCAACCTTGGCGCCCAATAGC 1082
Db 1886 GGAATACAGCTGTGCGCGCTCTCACTGGTGAAGAAAAACCAACCTTGGCGCCCAATAGC 1945
Qy 1083 CAAACCGCCTCTCCCGCGCGTGTGGCCGATTCAATTAATGCAGCTGGCAGCAGGTTTCC 1142
Db 1946 CAAACCGCCTCTCCCGCGCGTGTGGCCGATTCAATTAATGCAGCTGGCAGCAGGTTTCC 2005
Qy 1143 CGACTGGAAGCGGCGAGTGAGCGCAACGAATTAATGTAGTTAGCTCACTCATTAGGC 1202
Db 2006 CGACTGGAAGCGGCGAGTGAGCGCAACGAATTAATGTAGTTAG--CGCGAATTGATC 2063
Qy 1203 ACAATTTCTCATGTTTGACAGCTTATCATGCTGACAGTGCACCAATGCTTCTGGCGTC 1262
Db 2064 TGAATTTCTCATGTTTGACAGCTTATCATGCTGACAGTGCACCAATGCTTCTGGCGTC 2123
Qy 1263 AGCGAGCCATCGGAAGCTGTGGTATGGCTGTGCAGTCTGTAATCACTGCATTAATTCGTG 1322
Db 2124 AGCGAGCCATCGGAAGCTGTGGTATGGCTGTGCAGTCTGTAATCACTGCATTAATTCGTG 2183
Qy 1323 TCGCTCAAGGCGCACTCCCGTTCTGGATAAATGTTTTTGGCGCGACATCAATACGGTTCT 1382
```


Db 2184 TCGCTCAAGGCCACTCCCGTTCGTGATYATGTTTTTTCGCCGCAATCATTAACGGTTCT 2243
QY 1383 GGCAAAATATCT 1394
Db 2244 GGCAAAATATCT 2255
RESULT 7
US-08-778-717-5
; Sequence 5, Application US/08778717
; Patent No. 6602689
; GENERAL INFORMATION:
; APPLICANT: UENO, EIICHI
; APPLICANT: NOBUYUKI, FUJII
; APPLICANT: OKADA, MASAHISA
; TITLE OF INVENTION: FUSED DNA SEQUENCE, FUSED PROTEIN
; TITLE OF INVENTION: EXPRESSED FROM SAID FUSED DNA SEQUENCE AND METHOD FOR
; TITLE OF INVENTION: EXPRESSING SAID DNA SEQUENCE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSER: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; ADDRESSER: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/778,717
; FILING DATE: 12-DEC-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 352225/1995
; FILING DATE: 28-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2084-031-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4557 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; ORIGINAL SOURCE:
; ORGANISM: E. COLI
; STRAIN: BL21 (DE3)
; PUBLICATION INFORMATION:
; AUTHORS: NOBUYUKI FUJII ET AL,
; TITLE: FUSED DNA SEQUENCE, FUSED PROTEIN EXPRESSED
; TITLE: FROM SAID FUSED DNA SEQUENCE AND METHOD OF
; TITLE: EXPRESSING SAID FUSED PROTEIN
; RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 TO 4557
US-08-778-717-5
Query Match 36.5%; Score 1204; DB 4; Length 4557;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCGACACCATCAATGGTCAAAACCTTTCGGGGTATGCATGATAGCGCCGGAAGAGA 60
Db 2563 CCGACACCATCAATGGTCAAAACCTTTCGGGGTATGCATGATAGCGCCGGAAGAGA 2622

QY 61 GTCAATTCAGGGTGGTGAATGTGAACCAAGTAAAGTTATACGATGTGCGAGAGTATCGG 120
Db 2623 GTCAATTCAGGGTGGTGAATGTGAACCAAGTAAAGTTATACGATGTGCGAGAGTATCGG 2682
QY 121 GTGCTCTTATCAGACCGTTTCCCGCGTGTGAACCAAGCGCAGCCACGCTTCTCGCAAAA 180
Db 2683 GTGCTCTTATCAGACCGTTTCCCGCGTGTGAACCAAGCGCAGCCACGCTTCTCGCAAAA 2742
QY 181 CGCGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTAATTCCTCAACCGCTGGCAC 240
Db 2743 CGCGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTAATTCCTCAACCGCTGGCAC 2802
QY 241 AACAACTGGCGGCAAAACAGTCTGTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 2803 AACAACTGGCGGCAAAACAGTCTGTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGC 2862
QY 301 ACGGCCCTCGCAAAATTCGCGGCGATTAATCTCGGCGCGATCAATCTGGGTGCGACGG 360
Db 2863 ACGGCCCTCGCAAAATTCGCGGCGATTAATCTCGGCGCGATCAATCTGGGTGCGACGG 2922
QY 361 TGGTGGTCTGATGGTAGAACGAGCGCGTGAAGCCTGTAAAGCGCGGTGCAAAATC 420
Db 2923 TGGTGGTCTGATGGTAGAACGAGCGCGTGAAGCCTGTAAAGCGCGGTGCAAAATC 2982
QY 421 TTCTCGCAACCGGTGAGTGGGCTGATTAATCTCGCTGGATGACCAAGGATGCA 480
Db 2983 TTCTCGCAACCGGTGAGTGGGCTGATTAATCTCGCTGGATGACCAAGGATGCA 3042
QY 481 TTGCTGTGGAAGTGCCTGCACTAATGTTCCGGCGTTATTTCTGATGTCTCTGACCA 540
Db 3043 TTGCTGTGGAAGTGCCTGCACTAATGTTCCGGCGTTATTTCTGATGTCTCTGACCA 3102
QY 541 CACCCATCAACAGTATTAATTTCTCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db 3103 CACCCATCAACAGTATTAATTTCTCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 3162
QY 601 TGGTGCATTTGGGTACCAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTGG 660
Db 3163 TGGTGCATTTGGGTACCAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTGG 3222
QY 661 CGCGTCTCGCTGCGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 3223 CGCGTCTCGCTGCGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 3282
QY 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 3283 CGGAACGGGAAGGCGACTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 3342
QY 781 ATGAGGGCATCGTTCCCACTGGGATGCTGGTTGCCAAGCATGAGTGGCGTGGCGCA 840
Db 3343 ATGAGGGCATCGTTCCCACTGGGATGCTGGTTGCCAAGCATGAGTGGCGTGGCGCA 3402
QY 841 TGGCGCCATTACGAGTCCGGGCTCGCGTGTGGTGGGATATCTCGTAGTGGGATAG 900
Db 3403 TGGCGCCATTACGAGTCCGGGCTCGCGTGTGGTGGGATATCTCGTAGTGGGATAG 3462
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAAACCAACCATCAAAAGGATTTTC 960
Db 3463 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAAACCAACCATCAAAAGGATTTTC 3522
QY 961 GCCTCTGGGGCAACACAGCGTGGACCGCTTGTGCTCAACTCTCTCAGGCGCAGGCGGTGA 1020
Db 3523 GCCTCTGGGGCAACACAGCGTGGACCGCTTGTGCTCAACTCTCTCAGGCGCAGGCGGTGA 3582
QY 1021 AGGCAATCAGTGTGGCGCTCTCACTGTGGAAGAAACCAACCTTGGCGCCCAATA 1080
Db 3583 AGGCAATCAGTGTGGCGCTCTCACTGTGGAAGAAACCAACCTTGGCGCCCAATA 3642
QY 1081 CGCAAAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAATGAGCTGGCAGCAGAGTTT 1140
Db 3643 CGCAAAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAATGAGCTGGCAGCAGAGTTT 3702
QY 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCACTTAG 1200

Db 3703 CCGACTGGAAAGCGGCGAGTGGCGAACGAATTAATGTAGTACTCATTAG 3762
Qy 1201 GCAC 1204
Db 3763 GCAC 3766

RESULT 8
US-08-148-675A-2
; Sequence 2, Application US/08148675A
; Patent No. 5506121
; GENERAL INFORMATION:
; APPLICANT: Skerra, Arne; Schmidt, Thomas
; TITLE OF INVENTION: FUSION PEPTIDES WITH BINDING ACTIVITY FOR
; TITLE OF INVENTION: STREPTAVIDIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/148,675A
; FILING DATE: 3-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 42 37 113.9
; FILING DATE: 03-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsai, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: HUBR 1041
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3632 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-148-675A-2

Query Match 36.4%; Score 1200.8; DB 1; Length 3832;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1202; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGAGAGA 60
Db 3 CCGACACCATCGAATGGCGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGAGAGA 62

Qy 61 GTCAATTACGGTGGTGAATGTGAACACAGTAACGTTTATACGATGTCGCAGAGTATGCCG 120
Db 63 GTCAATTACGGTGGTGAATGTGAACACAGTAACGTTTATACGATGTCGCAGAGTATGCCG 122

Qy 121 GTGTCCTCTATCAGACCGTTTCCCGCGTGGTGAACACAGTAACGTTTATACGATGTCGCAGAGTATGCCG 120
Db 123 GTGTCCTCTATCAGACCGTTTCCCGCGTGGTGAACACAGTAACGTTTATACGATGTCGCAGAGTATGCCG 122

Qy 181 CGCGGGAAAAAGTGAAGCGCGCATGGCGAGCTGAATTACATTCCCAACCGCGTGGCAC 240
Db 183 CGCGGGAAAAAGTGAAGCGCGCATGGCGAGCTGAATTACATTCCCAACCGCGTGGCAC 242

Qy 241 AACNACTGGCGGGCAACAGTGGTGTGATGTGGCGTGGCACCTCAGTCTGGCCCTGC 300

Db 243 AACAACTGGCGGCAACAGATCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 302
Qy 301 ACGCGCGTTCGCAAAATTTGTGCGCGGATTAATCTCGCGCGGATCAATCTGGGTGCCACG 360
Db 303 ACGCGCGTTCGCAAAATTTGTGCGCGGATTAATCTCGCGCGGATCAATCTGGGTGCCACG 362
Qy 361 TGGTGGTGTGATGTAGAACGCGCTGTAAGCGCTGTAAGCGCGGATTAATCTGGGTGCCACATC 420
Db 363 TGGTGGTGTGATGTAGAACGCGCTGTAAGCGCGCTGTAAGCGCGGATTAATCTGGGTGCCACATC 422
Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATCTAATCTCCGCTGGATGACACAGATGCCA 480
Db 423 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATCTAATCTCCGCTGGATGACACAGATGCCA 482
Qy 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGGTTATTTCTTGATGTCTCTGACGAGA 540
Db 483 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGGTTATTTCTTGATGTCTCTGACGAGA 542
Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db 543 CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 602
Qy 601 TGGTGGCATTTGGGTCAACAGCAATCGCTGTTAGCGGCCCCATTAAGTTCTGTCTCGG 660
Db 603 TGGTGGCATTTGGGTCAACAGCAATCGCTGTTAGCGGCCCCATTAAGTTCTGTCTCGG 662
Qy 661 CGGCTCTGCGTCTGGCTGGCTGGCATATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 663 CGGCTCTGCGTCTGGCTGGCTGGCATATAATCTCACTCGCAATCAAAATTCAGCCGATAG 722
Qy 721 CGGAACGGGAAGCGCACTGGAGTGCCATGTCGGTTTCAACAAACCATGCAAAATGCTGA 780
Db 723 CGGAACGGGAAGCGCACTGGAGTGCCATGTCGGTTTCAACAAACCATGCAAAATGCTGA 782
Qy 781 ATGAGGCGATCTGTTCCCATGTCGATGCTGTTGCCAACCATGAGATGGCGTGGCGCAA 840
Db 783 ATGAGGCGATCTGTTCCCATGTCGATGCTGTTGCCAACCATGAGATGGCGTGGCGCAA 842
Qy 841 TGGCGCGCATTTACCGAGTCCGGCTCGCGGTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 843 TGGCGCGCATTTACCGAGTCCGGCTCGCGGTGGTGGGATATCTCGGTAGTGGGATACG 902
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAACAGGATTTTC 960
Db 903 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAACAGGATTTTC 962
Qy 961 GCCTGCTGGGCAACACAGGCTGGACCGCTTGTGTCAACTCTCTAGGGCGAGCGGTGA 1020
Db 963 GCCTGCTGGGCAACACAGGCTGGACCGCTTGTGTCAACTCTCTAGGGCGAGCGGTGA 1022
Qy 1021 AGGGCAATCAGCTGTTGCCCCGTCTCACTGGTGAAGAAAAAACCCCTGGCGCCCAATA 1080
Db 1023 AGGGCAATCAGCTGTTGCCCCGTCTCACTGGTGAAGAAAAAACCCCTGGCGCCCAATA 1082
Qy 1081 CGCAAAACCGCTCTCCCCCGGGTTGGCGGATTCATTAAATGAGCTGGCACGACAGGTTT 1140
Db 1083 CGCAAAACCGCTCTCCCCCGGGTTGGCGGATTCATTAAATGAGCTGGCACGACAGGTTT 1142
Qy 1141 CCGCTGTGAAGCGGCGAGTGGCGCAACGCAATTAATGTAGTTAGTCTACTATTAG 1200
Db 1143 CCGCTGTGAAGCGGCGAGTGGCGCAACGCAATTAATGTAGTTAGTCTACTATTAG 1202
Qy 1201 GCAC 1204
Db 1203 GCAC 1206

RESULT 9
US-08-487-283A-18/c
; Sequence 18, Application US/08487283A
; Patent No. 6355245
; GENERAL INFORMATION:
; APPLICANT: Evans, Mark J.

```

; APPLICANT: Matis, Louis A.
; APPLICANT: Mueller, Eileen Elliott
; APPLICANT: Nye, Steven H.
; APPLICANT: Rollins, Scott
; APPLICANT: Rother, Russell P.
; APPLICANT: Springhorn, Jeremy P.
; APPLICANT: Squinto, Stephen P.
; APPLICANT: Thomas, Thomas C.
; APPLICANT: Wilkins, James A.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT
; OF INFLAMMATORY DISEASES
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seth A. Fidel
; STREET: 25 Science Park (Alexion)
; CITY: New Haven
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06511
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.4mb storage
; COMPUTER: Macintosh Cetris 610
; OPERATING SYSTEM: System 7
; SOFTWARE: WordPerfect 3.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,283A
; FILING DATE: June 7, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/236,208
; FILING DATE: 02-MAY-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Seth A. Fidel.
; REGISTRATION NUMBER: 38,449
; REFERENCE/DOCKET NUMBER: ALX-152.1 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (203)776-1790
; TELEFAX: (203)772-3655
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5248 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Circular
; MOLECULE TYPE: Other nucleic acid
; DESCRIPTION: pET Trc S05/N1
; DESCRIPTION: prokaryotic expression vector
; US-08-487-283A-18

```

Query Match	36.3%	Score 1199.2	DB 3	Length 5248
Best Local Similarity	99.8%	Pred. No. 0		
Matches 1201	Conservative 0	Mismatches 3	Indels 0	Gaps 0
Qy	1	CCGACACATCGATGTTGCAAAACCTTTCGCGGTATGGCATGATACGCCCGCGGAAGAGA	60	
Db	4810	CGGACACATCGAATGGCGCAAAACCTTTCGCGGTATGGCATGATACGCCCGCGGAAGAGA	4751	
Qy	61	GTCAATTTCAGGTCGGTGAATGTGAAACCAAGTAACGTTATACGATGTCGCAGAGTATGCCG	120	
Db	4750	GTCAATTTCAGGTCGGTGAATGTGAAACCAAGTAACGTTATACGATGTCGCAGAGTATGCCG	4691	
Qy	121	GTGTCTCTTATCAGACCGTTTCCGCGTGTGTGAACACGAGGCAGCCAGTTTCTGCGAAAA	180	
Db	4690	GTGTCTCTTATCAGACCGTTTCCGCGTGTGTGAACACGAGGCAGCCAGTTTCTGCGAAAA	4631	
Qy	181	CGCGGGAAGAAGTGGAAAGCGCGATGGGAGCTGAATTACATTCCCAACCGCGTGGCAC	240	
Db	4630	CGCGGGAAGAAGTGGAAAGCGCGATGGGAGCTGAATTACATTCCCAACCGCGTGGCAC	4571	
Qy	241	AACAACATGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTTGGCACCTCCAGTCTGGCCCTGC	300	
Db	4570	AACAACATGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTTGGCACCTCCAGTCTGGCCCTGC	4511	
Qy	301	ACGCGCGCTCGCAAAATTGTTCGCGCGGATTTAAATCTCGCGCGGATCAACTGGGTGCCACGC	360	

```

RESULT 10
PCT-US96-08611A-21/c
: Sequence 21, Application PC/TUS96050611A
: GENERAL INFORMATION:
: APPLICANT: Mueller, John P.
: APPLICANT: Leonardo, Michael J.
: APPLICANT: McFarland, Henry F.
: APPLICANT: Matis, Louis A.
: APPLICANT: Mueller, Eileen Elliott

```

4510	Db	ACGGCCGTCGCAAAATTGTCGGCGCGATTAAATCTCGCGCCGATCAACTGGGTGCCACGC	4451
361	Qy	TGTTGGTGTTCGATGGTAGAACGAAGCGGCGTGAAGCCTGTAAAGCGGCGGTGCACAATC	420
4450	Db	TGTTGGTGTTCGATGGTAGAACGAAGCGGCGTGAAGCCTGTAAAGCGGCGGTGCACAATC	4391
421	Qy	TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATATCCGCTGGATGACACAGGATGCCA	480
4390	Db	TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATATCCGCTGGATGACACAGGATGCCA	4331
481	Qy	TTGCTGTGGAAAGCTGCCTGCACATAATGTTCCGCGCTTATTTCTTTGATGTCTCTGACCAGA	540
4330	Db	TTGCTGTGGAAAGCTGCCTGCACATAATGTTCCGCGCTTATTTCTTTGATGTCTCTGACCAGA	4271
541	Qy	CACCCATCAACAGTATATTTTCTCCCATGAAGA CGGTACGCGACTGGGCGTGGAGCATC	600
4270	Db	CACCCATCAACAGTATATTTTCTCCCATGAAGA CGGTACGCGACTGGGCGTGGAGCATC	4211
601	Qy	TGCTCGCATTTGGGCTACCCAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG	660
4210	Db	TGCTCGCATTTGGGCTACCCAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG	4151
661	Qy	CGGCTCTGCGTCTGGCTGGCTGGCATAAAATATCTCACTCGCAATCAAAATTCAGCGGATAG	720
4150	Db	CGGCTCTGCGTCTGGCTGGCTGGCATAAAATATCTCACTCGCAATCAAAATTCAGCGGATAG	4091
721	Qy	CGGAAACGGGAAGCGGACTGAGGTGCCATGTCGCGTTTTTCAACAAACCATGCAAAATGCTGA	780
4090	Db	CGGAAACGGGAAGCGGACTGAGGTGCCATGTCGCGTTTTTCAACAAACCATGCAAAATGCTGA	4031
781	Qy	ATGAGGGCATCGTTTCCCACTCGCATGCTGCTGCTGCCAACGATCAGATCGCGTGGCGCGCA	840
4030	Db	ATGAGGGCATCGTTTCCCACTCGCATGCTGCTGCTGCCAACGATCAGATCGCGTGGCGCGCA	3971
841	Qy	TGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGGCGATATCTCGGTAGTGGGATACG	900
3970	Db	TGCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGGCGATATCTCGGTAGTGGGATACG	3911
901	Qy	ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAACCCATCAAAACAGGATTTTC	960
3910	Db	ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAACCCATCAAAACAGGATTTTC	3851
961	Qy	GCCTGCTGGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA	1020
3850	Db	GCCTGCTGGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA	3791
1021	Qy	AGGGCAATCAGCTGTTTCCCGCTCTCACTGGTGAAGAAAAACCACTCGCGCCCCAATA	1080
3790	Db	AGGGCAATCAGCTGTTTCCCGCTCTCACTGGTGAAGAAAAACCACTCGCGCCCCAATA	3731
1081	Qy	CGCAACCGGCTCTCCCGCGGCTTGGCCGATTCATTAAATGCAGCTGGCAGCAGGTTT	1140
3730	Db	CGCAACCGGCTCTCCCGCGGCTTGGCCGATTCATTAAATGCAGCTGGCAGCAGGTTT	3671
1141	Qy	CCGACTTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200
3670	Db	CCGACTTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	3611
1201	Qy	GCAC 1204	
3610	Db	GCAC 3607	

APPLICANT: Nye, Steven H.
 APPLICANT: Pelfrey, Clara M.
 APPLICANT: Squinto, Stephen P.
 APPLICANT: Wilkins, James A.
 TITLE OF INVENTION: Modified Myelin Protein Molecules
 NUMBER OF SEQUENCES: 29
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Maurice M. Klee
 STREET: 1951 Burr Street
 CITY: Fairfield
 STATE: Connecticut
 COUNTRY: USA
 ZIP: 06430
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5 inch, 0.8 Mb storage
 COMPUTER: Macintosh Centris 610
 OPERATING SYSTEM: System 7
 SOFTWARE: Microsoft Word 6.0.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US96/05611A
 FILING DATE: 02-MAY-1995
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/431,644
 FILING DATE: May 2, 1995
 APPLICATION NUMBER: 08/431,648
 FILING DATE: May 2, 1995
 APPLICATION NUMBER: 08/482,114
 FILING DATE: June 7, 1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Klee, Maurice M.
 REGISTRATION NUMBER: 30,399
 REFERENCE/DOCKET NUMBER: ALX-129
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (203) 255 1400
 TELEFAX: (203) 254 1101
 INFORMATION FOR SEQ ID NO: 21:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5248 base pairs
 TYPE: Nucleic Acid
 STRANDEDNESS: Double
 TOPOLOGY: Circular
 MOLECULE TYPE: Other nucleic acid
 DESCRIPTION: PET Trc S05/NI
 DESCRIPTION: prokaryotic expression vector
 PCT-US96-05611A-21

Query Match 36.3%; Score 1199.2; DB 5; Length 5248;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 CCGACACCATCGAATGGTGAACACCTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
 DB 4810 CCGACACCATCGAATGGTGAACACCTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA 4751
 QY 61 GTCAATTGAGGTGGTGAATGTAAGTGAACACCTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA 120
 DB 4750 GTCAATTGAGGTGGTGAATGTAAGTGAACACCTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA 4691
 QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACCGCGAGCGACGCTTTCGCGAATA 180
 DB 4690 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACCGCGAGCGACGCTTTCGCGAATA 4631
 QY 181 CCGGGGAAAAGTGAACCGCGATGCGGAGCTGAATTACATTCCTCCAAACCGCGTGGCAC 240
 DB 4630 CCGGGGAAAAGTGAACCGCGATGCGGAGCTGAATTACATTCCTCCAAACCGCGTGGCAC 4571
 QY 241 AACAACTGGCGGGCAACAGTGGTGTGATGGCGTTCACCTCCAGTCTGGCCCTGC 300
 DB 4570 AACAACTGGCGGGCAACAGTGGTGTGATGGCGTTCACCTCCAGTCTGGCCCTGC 4511
 QY 301 ACGCGCGTGCAGAAATGTTCGGCGGATTAATCTCCGCGCGATCACTGGGTGCGACGCG 360

DB 4510 ACGCGCGTGCAGAAATGTTCGGCGGATTAATCTCCGCGCGATCAACTGGTGGCGACG 4451
 QY 361 TGGTGGTGTGATGATGTAAGAGCGCGTTCGAAGCGCTGTAAAGCGCGGTGCACAAATC 420
 DB 4450 TGGTGGTGTGATGATGTAAGAGCGCGTTCGAAGCGCTGTAAAGCGCGGTGCACAAATC 4391
 QY 421 TTCTTCGGCGCAACCGCTCAGTGGCGTATCACTTAACCTATCCGCTGGATGACCGAGATGCCA 480
 DB 4390 TTCTTCGGCGCAACCGCTCAGTGGCGTATCACTTAACCTATCCGCTGGATGACCGAGATGCCA 4331
 QY 481 TTGCTGTGGAAGCTGCCTGCACATAATGTTCGGCGTATTAATCTCTGATGTCTCTACCCAGA 540
 DB 4330 TTGCTGTGGAAGCTGCCTGCACATAATGTTCGGCGTATTAATCTCTGATGTCTCTACCCAGA 4271
 QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGCGTGGAGATC 600
 DB 4270 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGCGTGGAGATC 4211
 QY 601 TGGTTCGCATTTGGGTACCCAGCAATCGCTGTAGCGGCGCAATTAAGTTCCTCTCGG 660
 DB 4210 TGGTTCGCATTTGGGTACCCAGCAATCGCTGTAGCGGCGCAATTAAGTTCCTCTCGG 4151
 QY 661 CGGTTCGCTCTGGGTGGCTGCGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
 DB 4150 CGGTTCGCTCTGGGTGGCTGCGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 4091
 QY 721 CGGAACGGGAAGCGGACGTGGAGTGCCATGTCCGGTTTTTCAACCAACGATGCAAAATGCTGA 780
 DB 4090 CGGAACGGGAAGCGGACGTGGAGTGCCATGTCCGGTTTTTCAACCAACGATGCAAAATGCTGA 4031
 QY 781 ATGAGGCGCATCGTTCCCATCTCGATGCTGGTTCGCCAACGATCAGATGGCGCTGGCGCAA 840
 DB 4030 ATGAGGCGCATCGTTCCCATCTCGATGCTGGTTCGCCAACGATCAGATGGCGCTGGCGCAA 3971
 QY 841 TGGCGGCGATTCAGGAGTCCGGCTGGCGCTGGTGGGATATCTCGGTAGTGGGATACG 900
 DB 3970 TGGCGGCGATTCAGGAGTCCGGCTGGCGCTGGTGGGATATCTCGGTAGTGGGATACG 3911
 QY 901 ACGATACCGAAGACAGCTCATGTTATCCGCGCTTAACCCATCAACACAGGATTTTC 960
 DB 3910 ACGATACCGAAGACAGCTCATGTTATCCGCGCTTAACCCATCAACACAGGATTTTC 3851
 QY 961 GCCTGTGGGGCAACACCGGTGGACCGCTTGTCGCAACTCTCTCAGGGCCAGCGGTGA 1020
 DB 3850 GCCTGTGGGGCAACACCGGTGGACCGCTTGTCGCAACTCTCTCAGGGCCAGCGGTGA 3791
 QY 1021 AGGGCAATCAGCTGTTCGCGCTCTCACTGGTGAAGAAAAACCCCTGGCGGCCCAATA 1080
 DB 3790 AGGGCAATCAGCTGTTCGCGCTCTCACTGGTGAAGAAAAACCCCTGGCGGCCCAATA 3731
 QY 1081 CGAAACCCGCTCTCCCGCGGTTGGCGGATTTCAATATGAGCTGGCAGCAGGTTT 1140
 DB 3730 CGAAACCCGCTCTCCCGCGGTTGGCGGATTTCAATATGAGCTGGCAGCAGGTTT 3671
 QY 1141 CCGCACTGAAAAGCGGCGAGTGAAGCGCAACGCAATTAATGTGAGTTAGTCTCACTCATTAG 1200
 DB 3670 CCGCACTGAAAAGCGGCGAGTGAAGCGCAACGCAATTAATGTGAGTTAGTCTCACTCATTAG 3611
 QY 1201 GCAC 1204
 DB 3610 GCAC 3607

RESULT 11
 US-10-263-103-35/c
 ; Sequence 35, Application US/10263103
 ; Patent No. 6821723
 ; GENERAL INFORMATION:
 ; APPLICANT: AVENTIS PASTEUR
 ; APPLICANT: Chevallier, Michel
 ; APPLICANT: El Habib, Raphaelle
 ; APPLICANT: Krell, Tino
 ; APPLICANT: Sodoyer, Regis

; TITLE OF INVENTION: gp41 antigen
; FILE REFERENCE: 01-1692-A
; CURRENT APPLICATION NUMBER: US/10/263,103
; CURRENT FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 5312
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: plasmid pET-cer
US-10-263-103-35

Query Match 36.3%; Score 1199.2; DB 4; Length 5312;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTCGGGTATGGCATGATAGCCCGGAAGAGA 60
DB 4622 CGGACACCATCGAATGGCGCAAAACCTTTCGGGTATGGCATGATAGCCCGGAAGAGA 4563

QY 61 GTCATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGACAGTATGCCG 120
DB 4562 GTCATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGACAGTATGCCG 4503

QY 121 GTGCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCACGTTTCTGCGAAAA 180
DB 4502 GTGCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCACGTTTCTGCGAAAA 4443

QY 181 CGCGGAAAAAGTGAAGCGGGGATGGCGGAGCTGAATTAATCCCAACCGCGTGGCAC 240
DB 4442 CGCGGAAAAAGTGAAGCGGGGATGGCGGAGCTGAATTAATCCCAACCGCGTGGCAC 4383

QY 241 AACCACTCGCGGCAAAACAGTCTGTTGCTGATTGGCGTTTGCACCTCCAGTCTGGCCCTGC 300
DB 4382 AACCACTCGCGGCAAAACAGTCTGTTGCTGATTGGCGTTTGCACCTCCAGTCTGGCCCTGC 4323

QY 301 ACGCGCGTTCGCAAAATTGTTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
DB 4322 ACGCGCGTTCGCAAAATTGTTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 4263

QY 361 TGGTGGTTCGATGTAGAAAGCAAGCGGCTGGAAGCGCTGTAAGCGGGGTGCAATTC 420
DB 4262 TGGTGGTTCGATGTAGAAAGCAAGCGGCTGGAAGCGCTGTAAGCGGGGTGCAATTC 4203

QY 421 TTCTCGCGCAACGCTGAGTGGGCTGATCAATTAATCTCCGCTGATGACACGAGGATGCA 480
DB 4202 TTCTCGCGCAACGCTGAGTGGGCTGATCAATTAATCTCCGCTGATGACACGAGGATGCA 4143

QY 481 TTGCTGTGGAAGCTGCCTGCATTAATGTTTCGGCGTATTTCTTGATGTCTTGACCCAGA 540
DB 4142 TTGCTGTGGAAGCTGCCTGCATTAATGTTTCGGCGTATTTCTTGATGTCTTGACCCAGA 4083

QY 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGATGGCGGTGGAGCATC 600
DB 4082 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGATGGCGGTGGAGCATC 4023

QY 601 TGGTCCGATTTGGTCACGACGAAATCGCGCTTTAGCGGGCCCATTAAGTTCTGTCGG 660
DB 4022 TGGTCCGATTTGGTCACGACGAAATCGCGCTTTAGCGGGCCCATTAAGTTCTGTCGG 3963

QY 661 CCGCTCTGCGTCTGGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 3962 CCGCTCTGCGTCTGGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 3903

QY 721 CGGAAACGGGAAGCGACATGGAGTGCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
DB 3902 CGGAAACGGGAAGCGACATGGAGTGCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 3843

QY 781 ATGAGGGCATCTTCCCACTGGATGCTGGTTGCCAACGATCAGATGGCGTGGCGGCA 840
DB 3842 ATGAGGGCATCTTCCCACTGGATGCTGGTTGCCAACGATCAGATGGCGTGGCGGCA 3783

QY 841 TGCAGCGCATTTACCGAGTCCGGGCTGGCGTGGTGGGATATCTCGGTAGTGGATACG 900
DB 3782 TGCAGCGCATTTACCGAGTCCGGGCTGGCGTGGTGGGATATCTCGGTAGTGGATACG 3723

QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTGAACACCATCAAAACAGGATTTTC 960
DB 3722 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTGAACACCATCAAAACAGGATTTTC 3663

QY 961 GCCTGCTGGGCAAAACACGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
DB 3662 GCCTGCTGGGCAAAACACGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 3603

QY 1021 AGGCAATCAGCTGTTGCCCTCTCACTGGTGAAGAAAAACACCTCGGGCCCAATA 1080
DB 3602 AGGCAATCAGCTGTTGCCCTCTCACTGGTGAAGAAAAACACCTCGGGCCCAATA 3543

QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGCAGCTGCGACAGGTTT 1140
DB 3542 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGCAGCTGCGACAGGTTT 3483

QY 1141 CCCGACTGGAAGCGGCGAGTGCAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
DB 3482 CCCGACTGGAAGCGGCGAGTGCAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTTAG 3423

QY 1201 GCAC 1204
DB 3422 GCAC 3419

RESULT 12
US-08-929-967-1/c
; Patent 1, Application US/08929967
; Patent No. 5891637
; GENERAL INFORMATION:
; APPLICANT: Ruppert, Siegfried J.W.
; TITLE OF INVENTION: Construction of Full-length cDNA Libraries
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/929,967
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1035R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5443 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-929-967-1

Query Match 36.3%; Score 1199.2; DB 2; Length 5443;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CGCACCATCGAATGGTGCAGAAACCTTTTCGGGGTATGCGATGATAGCGCCCGGAGAGA 60
Db 4810 CGGACACCATCGAATGGTGCAGAAACCTTTTCGGGGTATGCGATGATAGCGCCCGGAGAGA 4751
QY 61 GTCAAATTCAGGGTGGTGAATGTGAACACAGTAAACAGTATACGATGTCGAGAGTATGCCG 120
Db 4750 GTCAAATTCAGGGTGGTGAATGTGAACACAGTAAACAGTATACGATGTCGAGAGTATGCCG 4691
QY 121 GTGTCTCTTATCAGACCGGTTTCGGGGTGTGAACACAGCGGACGACGCTTTCTGCGAAAA 180
Db 4690 GTGTCTCTTATCAGACCGGTTTCGGGGTGTGAACACAGCGGACGACGCTTTCTGCGAAAA 4631
QY 181 CGCGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 240
Db 4630 CGCGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 4571
QY 241 AACAACTGGCGGCAACAGTGGTGTGATTTGGGGTTCGCACTTCCAGTCTGGCCCTGC 300
Db 4570 AACAACTGGCGGCAACAGTGGTGTGATTTGGGGTTCGCACTTCCAGTCTGGCCCTGC 4511
QY 301 ACGCGCGTTCGCAAAATTTGCGGGCGATTAATCTCGCGCGATCAACTCGCGTGGCGGCGG 360
Db 4510 ACGCGCGTTCGCAAAATTTGCGGGCGATTAATCTCGCGCGATCAACTCGCGTGGCGGCGG 4451
QY 361 TGGTGGTTCGATGTGAGAACGAGCGGCGTGAAGCGGCTGAAGCGGCGGTCGACAAATC 420
Db 4450 TGGTGGTTCGATGTGAGAACGAGCGGCGTGAAGCGGCTGAAGCGGCGGTCGACAAATC 4391
QY 421 TTCTCGCGCAACGCGTCACTGGGCTGATTAATCTCGCGCGATCAACTCGCGTGGATGCGCA 480
Db 4390 TTCTCGCGCAACGCGTCACTGGGCTGATTAATCTCGCGCGATCAACTCGCGTGGATGCGCA 4331
QY 481 TTGCTGTGGAAGCTGCCTGCATTAATGTTCCGGCGTATTTCTTGATGTTCTCTGACGAGA 540
Db 4330 TTGCTGTGGAAGCTGCCTGCATTAATGTTCCGGCGTATTTCTTGATGTTCTCTGACGAGA 4271
QY 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGGATCGGCGTGGAGCATC 600
Db 4270 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGGATCGGCGTGGAGCATC 4211
QY 601 TGGTGCATTTGGGTCACGCAAAATCGCGCTGTAGCGGGCCATTAAGTCTGCTCTCGG 660
Db 4210 TGGTGCATTTGGGTCACGCAAAATCGCGCTGTAGCGGGCCATTAAGTCTGCTCTCGG 4151
QY 661 CGCGTCTGCGTCTGGCTGGGATAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db 4150 CGCGTCTGCGTCTGGCTGGGATAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 4091
QY 721 CGGAAACGGGAAGGCGACTGGAGTCCCATGTCGGGTTTCAACAAACCATGCAATGCTGA 780
Db 4090 CGGAAACGGGAAGGCGACTGGAGTCCCATGTCGGGTTTCAACAAACCATGCAATGCTGA 4031
QY 781 ATGAGGGCATCGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
Db 4030 ATGAGGGCATCGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGGCGCTGGGCGCAA 3971
QY 841 TGGCGGCATTAACGAGTCCGGGTCGGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
Db 3970 TGGCGGCATTAACGAGTCCGGGTCGGGTTGGTGGGATATCTCGGTAGTGGGATAG 3911
QY 901 ACGATACCGAAGACAGCTCATGTTTATATCCGCGGTTTAAACCATCAACAGGATTTTC 960
Db 3910 ACGATACCGAAGACAGCTCATGTTTATATCCGCGGTTTAAACCATCAACAGGATTTTC 3851
QY 961 GCTGTGGGGCAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 3850 GCTGTGGGGCAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 3791
QY 1021 AGGCAATCAGCTGTTGCCGCTCTCACTGTTGAAGAAAAACCACTTGGCGCCCAATA 1080
Db 3790 AGGCAATCAGCTGTTGCCGCTCTCACTGTTGAAGAAAAACCACTTGGCGCCCAATA 3731
QY 1081 CGCAAAACCGCTCTCTCCCGCGCGTGGCCGATTAATTAATGACGCTGGGACGACGAGTTT 1140

Db 3730 CGCAAAACCGCTCTCTCCCGCGCGTGGCCGATTAATTAATGACGCTGGCAGCAGGTTT 3671
QY 1141 CCCCACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTGTAGCTCACTCATTAG 1200
Db 3670 CCCCACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTGTAGCTCACTCATTAG 3611
QY 1201 GCAC 1204
Db 3610 GCAC 3607

RESULT 13
US-09-702-705-785/c
; Sequence 785, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-785

Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTTCGGGGTATGCGATGATAGCGCCCGGAGAGA 60
Db 4677 CGGACACCATCGAATGGTGCAGAAACCTTTTCGGGGTATGCGATGATAGCGCCCGGAGAGA 4618
QY 61 GTCAAATTCAGGGTGGTGAATGTGAACACAGTAAACAGTATACGATGTCGAGAGTATGCCG 120
Db 4617 GTCAAATTCAGGGTGGTGAATGTGAACACAGTAAACAGTATACGATGTCGAGAGTATGCCG 4558
QY 121 GTGTCTCTTATCAGACCGGTTTCGGGGTGTGAACACAGCGGACGCTTTCTGCGAAAA 180
Db 4557 GTGTCTCTTATCAGACCGGTTTCGGGGTGTGAACACAGCGGACGCTTTCTGCGAAAA 4498
QY 181 CGCGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 240
Db 4497 CGCGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 4438
QY 241 AACAACTGGCGGCAACAGTGGTGTGATTTGGGGTTCGCACTTCCAGTCTGGCCCTGC 300
Db 4437 AACAACTGGCGGCAACAGTGGTGTGATTTGGGGTTCGCACTTCCAGTCTGGCCCTGC 4378
QY 301 ACGCGCGTTCGCAAAATTTGCGGGCGATTAATCTCGCGCGATCAACTCGGTCGACGG 360
Db 4377 ACGCGCGTTCGCAAAATTTGCGGGCGATTAATCTCGCGCGATCAACTCGGTCGACGG 4318
QY 361 TGGTGGTTCGATGTGAGAACGAGCGGCGTGAAGCGGCTGTAAAGCGGCGGTCGCAATC 420
Db 4317 TGGTGGTTCGATGTGAGAACGAGCGGCGTGAAGCGGCTGTAAAGCGGCGGTCGCAATC 4258
QY 421 TTCTCGCGCAACGCGTCACTGGGCTGATTAATTAATGACGCTGGGACGACGAGTTT 480

Db 4257 TTCTCGCAACGCGTCAGTGGCTGATCATTAATACTATCCGCTGGATGACCAAGATGCCA 4198
QY 481 TTGCTGTGGAAGCTGCCTGCAATAAGTTCGGCGGTATTTCTTGATGTCCTGACCAAGA 540
Db 4197 TTGCTGTGGAAGCTGCCTGCAATAAGTTCGGCGGTATTTCTTGATGTCCTGACCAAGA 4138
QY 541 CACCCATCAACAGTATATTTTCTCCCATGAAGCGGTACGCACTGGCGGTGGAGCATC 600
Db 4137 CACCCATCAACAGTATATTTTCTCCCATGAAGCGGTACGCACTGGCGGTGGAGCATC 4078
QY 601 TGGTGCATTTGGTTCACCAAGCAAAATCGCGCTGTTAGCGGGGCCCAATAAGTTCCTGCTCGG 660
Db 4077 TGGTGCATTTGGTTCACCAAGCAAAATCGCGCTGTTAGCGGGGCCCAATAAGTTCCTGCTCGG 4018
QY 661 CGCGTCTGCGTCTGCGTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4017 CGCGTCTGCGTCTGCGTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 3958
QY 721 CGGAACGGGAAGGCGACATGGAGTGCATGTCGGGTTTTCACAAACCAATGCAATGCTGA 780
Db 3957 CGGAACGGGAAGGCGACATGGAGTGCATGTCGGGTTTTCACAAACCAATGCAATGCTGA 3898
QY 781 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTGCCAAGATGCGCTGGCGCAA 840
Db 3897 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTGCCAAGATGCGCTGGCGCAA 3838
QY 841 TCGCGCCATTAACGAGTCCGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATAG 900
Db 3837 TCGCGCCATTAACGAGTCCGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATAG 3778
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTAAACCAACCAATCAAAACGAGATTTTC 960
Db 3777 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTAAACCAACCAATCAAAACGAGATTTTC 3718
QY 961 GCCTCTGGGGCAACACGCGTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 3717 GCCTCTGGGGCAACACGCGTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 3658
QY 1021 AGGCAATCAGCTGTGGCGGTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCCAATA 1080
Db 3657 AGGCAATCAGCTGTGGCGGTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCCAATA 3598
QY 1081 CGCAACCGCTCTCCCGCGGTGGCGGATTCATTAATGCACTGGCAGCAGCAAGGTTT 1140
Db 3597 CGCAACCGCTCTCCCGCGGTGGCGGATTCATTAATGCACTGGCAGCAGCAAGGTTT 3538
QY 1141 CCGCACTGGAAGCGGCGAGTCAGCGCAACGCAATTAATGTCAGTTAGTCACTCATTTAG 1200
Db 3537 CCGCACTGGAAGCGGCGAGTCAGCGCAACGCAATTAATGTCAGTTAGTCACTCATTTAG 3478
QY 1201 GCAC 1204
Db 3477 GCAC 3474

RESULT 14

US-09-736-457-785/c
; Sequence 785, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Panger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15

; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-785

Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGGCAAAACCTTTTCGGGTATGGCATGATAGCCCGGAAGAGA 60
Db 4677 CGGACACCATCGAATGGCGCAAAACCTTTTCGGGTATGGCATGATAGCCCGGAAGAGA 4618
QY 61 GTCAATTCAGGGTGGTGAATGTAACAGTAACAGTTATACGATGTCGACAGTATGCCG 120
Db 4617 GTCAATTCAGGGTGGTGAATGTAACAGTAACAGTTATACGATGTCGACAGTATGCCG 4558
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACAGGCCAGCCACCGTTTCTGCGAAA 180
Db 4557 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACAGGCCAGCCACCGTTTCTGCGAAA 4498
QY 181 CGCGGAAAAAGTGGAGCGCGATGGCGAGCTGAATTAATTCCTCCCAACCGCGTGGCA 240
Db 4497 CGCGGAAAAAGTGGAGCGCGATGGCGAGCTGAATTAATTCCTCCCAACCGCGTGGCA 4438
QY 241 AACAACTGGCGGGCAACAGTCTGATGTTGGGTTGCCACCTCCAGTCTGGGCCCTGC 300
Db 4437 AACAACTGGCGGGCAACAGTCTGATGTTGGGTTGCCACCTCCAGTCTGGGCCCTGC 4378
QY 301 ACGGCGCTCGCAAAATTCGCGCGGATTAATTCGCGCGCATCAACTGGGTGGCAGCG 360
Db 4377 ACGGCGCTCGCAAAATTCGCGCGGATTAATTCGCGCGCATCAACTGGGTGGCAGCG 4318
QY 361 TGGTGGTTCGATGGTGAACAGAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCAATATC 420
Db 4317 TGGTGGTTCGATGGTGAACAGAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCAATATC 4258
QY 421 TTCTCGCGCAACGCGTCACTGGGTGATCAATTAATTCGCTGGATGACCAAGGATGCA 480
Db 4257 TTCTCGCGCAACGCGTCACTGGGTGATCAATTAATTCGCTGGATGACCAAGGATGCA 4198
QY 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGGTTATTTCTTGATGTCCTGACCAAGA 540
Db 4197 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGGTTATTTCTTGATGTCCTGACCAAGA 4138
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACCGCATCGGCGGTGGAGCATC 600
Db 4137 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACCGCATCGGCGGTGGAGCATC 4078
QY 601 TGGTGCATTTGGTTCACCAAGCAAAATCGCGCTGTTAGCGGGGCCCAATAAGTTCCTGCTCGG 660
Db 4077 TGGTGCATTTGGTTCACCAAGCAAAATCGCGCTGTTAGCGGGGCCCAATAAGTTCCTGCTCGG 4018
QY 661 CGCGTCTGCGTCTGCGTGGCGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4017 CGCGTCTGCGTCTGCGTGGCGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 3958
QY 721 CGGAACGGGAAGGCGACATGGAGTGCATGCTGGTGGCAACGATGCGCTGGCGCGCAA 840
Db 3957 CGGAACGGGAAGGCGACATGGAGTGCATGCTGGTGGCAACGATGCGCTGGCGCGCAA 3898
QY 781 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTGGCAACGATGCGCTGGCGCGCAA 840
Db 3897 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTGGCAACGATGCGCTGGCGCGCAA 3838
QY 841 TCGCGCCATTAACGAGTCCGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATAG 900
Db 3837 TCGCGCCATTAACGAGTCCGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATAG 3778


```
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACACCATCAAAACAGGATTTTC 960
Db 3777 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACACCATCAAAACAGGATTTTC 3718
QY 961 GCCTGCTGGGGCAAAACAGCGTGGACCGTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 3717 GCCTGCTGGGGCAAAACAGCGTGGACCGTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 3658
QY 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAACACCCCTGGCGCCCAATA 1080
Db 3657 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAACACCCCTGGCGCCCAATA 3598
QY 1081 CGCAACCGCTCTCCCGCGGTTGGCGAATTAATGAGCTGGCAGACAGGTTT 1140
Db 3597 CGCAACCGCTCTCCCGCGGTTGGCGAATTAATGAGCTGGCAGACAGGTTT 3538
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTACCTCACTCATTTAG 1200
Db 3537 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTACCTCACTCATTTAG 3478
QY 1201 GCAC 1204
Db 3477 GCAC 3474
```

RESULT 15

```
US-09-614-124B-785/c
; Sequence 785, Application US/09614124B
; Patent No. 6630574
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
```

```
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
```

```
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
```

```
; FILE REFERENCE: 210121.478C9
```

```
; CURRENT APPLICATION NUMBER: US/09/614.124B
```

```
; CURRENT FILING DATE: 2001-07-11
```

```
; NUMBER OF SEQ ID NOS: 1688
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 785
```

```
; LENGTH: 5502
```

```
; TYPE: DNA
```

```
; ORGANISM: Homo sapien
```

```
US-09-614-124B-785
```

```
Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 1 CCGACACCATCGATGGTGCAAAACCTTTCGCGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 4677 CGGACACCATCGATGGTGCAAAACCTTTCGCGTATGGCATGATAGCGCCCGGAAGAGA 4618
QY 61 GTCAATTCAGGGTGGTGAATGTGAACACAGTTATACGTTATACGATGTCGACAGATATGCCG 120
Db 4617 GTCAATTCAGGGTGGTGAATGTGAACACAGTTATACGTTATACGATGTCGACAGATATGCCG 4558
QY 121 GTGTCTTTATCAGACCGTTTCCCGTGGTGAACACAGGCGAGCCAGTTTCTGCGAATA 180
Db 4557 GTGTCTTTATCAGACCGTTTCCCGTGGTGAACACAGGCGAGCCAGTTTCTGCGAATA 4498
QY 181 CGCGGAAAAGTGAAGCGCGATGCGGAGCTGAATTACATTCCTCAACCGCGTGGCAC 240
Db 4497 CGCGGAAAAGTGAAGCGCGATGCGGAGCTGAATTACATTCCTCAACCGCGTGGCAC 4438
QY 241 AACAACTGGCGGCAACACAGTCGTTGCTGATGGCGTTGCCACCTCCAGCTCTGGCCCTGC 300
```

```
Search completed: August 3, 2005, 02:32:27
Job time : 357.4 secs
```

```
Db 4437 AACAACTGGCGGCAACACAGTCGTTGCTGATGGCGTTGCCACTCCAGTCTGGCCCTGC 4378
QY 301 ACGCGCGTCGCAAAATTTGTCCGCGCATTAATCTCCGCGCGATCAACTGGGTGCCACGC 360
Db 4377 ACGCGCGTCGCAAAATTTGTCCGCGCATTAATCTCCGCGCGATCAACTGGGTGCCACGC 4318
QY 361 TGGTGGTGTGATGGTGAAGCAAGCGCGTGCAGAGCTGTAAAGCGCGGTGCACAAATC 420
Db 4317 TGGTGGTGTGATGGTGAAGCAAGCGCGTGCAGAGCTGTAAAGCGCGGTGCACAAATC 4258
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACACAGATGCCA 480
Db 4257 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACACAGATGCCA 4198
QY 481 TTGCTGGGAAGCTGCTGCACATAATTTCCGCGTTAATTTCTTGATGTCTCTGACCAGA 540
Db 4197 TTGCTGGGAAGCTGCTGCACATAATTTCCGCGTTAATTTCTTGATGTCTCTGACCAGA 4138
QY 541 CACCATCAACAGCTAATTTTCTCCATGAAGCGGTACGCACTGGGCGTGGAGATC 600
Db 4137 CACCATCAACAGCTAATTTTCTCCATGAAGCGGTACGCACTGGGCGTGGAGATC 4078
QY 601 TGGTTCGCAATTTGGGTCAACAGCAAAATCGCGTGTAGCGGGCCCAATTAAGTTCTGTCGG 660
Db 4077 TGGTTCGCAATTTGGGTCAACAGCAAAATCGCGTGTAGCGGGCCCAATTAAGTTCTGTCGG 4018
QY 661 CGCGTCTGCGTCTGGCTGGCTGGCGATAATATCTCACTGCAATCAAAATTCAGCCGATAG 720
Db 4017 CGCGTCTGCGTCTGGCTGGCTGGCGATAATATCTCACTGCAATCAAAATTCAGCCGATAG 3958
QY 721 CGGAACGGGAAGCGCACTGGAGTGCCAATGTCGCGTTTCAACAAACCATGCAATGCTGA 780
Db 3957 CGGAACGGGAAGCGCACTGGAGTGCCAATGTCGCGTTTCAACAAACCATGCAATGCTGA 3898
QY 781 ATGAGGCAATCGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
Db 3897 ATGAGGCAATCGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGGCGCTGGGCGCAA 3838
QY 841 TGGCGGCCATTACCGAGTCGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 900
Db 3837 TGGCGGCCATTACCGAGTCGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 3778
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAACCATCAAAACAGGATTTTC 960
Db 3777 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAACCATCAAAACAGGATTTTC 3718
QY 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 3717 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 3658
QY 1021 AGGGCAATCAGCTGTTTGGCCGCTCTCACTGGTGAAGAAAACCAACCCCTGGCGCCCAATA 1080
Db 3657 AGGGCAATCAGCTGTTTGGCCGCTCTCACTGGTGAAGAAAACCAACCCCTGGCGCCCAATA 3598
QY 1081 CGCAAAACCCCTCTCCCGCGGTTTGGCGGATTCATTAATGAGCTGGCAGCAGCAGGTTT 1140
Db 3597 CGCAAAACCCCTCTCCCGCGGTTTGGCGGATTCATTAATGAGCTGGCAGCAGGTTT 3538
QY 1141 CCGGACTGAAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGTCTCACTATTAG 1200
Db 3537 CCGGACTGAAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGTCTCACTATTAG 3478
QY 1201 GCAC 1204
Db 3477 GCAC 3474
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 14:24:04 ; Search time 1094.6 Seconds
(without alignments)
17846.836 Million cell updates/sec

Title: US-09-765-555B-16
Perfect score: 3300
Sequence: 1 ccgacacatcgatgtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2002bs:*
8: Geneseqn2003as:*
9: Geneseqn2003bs:*
10: Geneseqn2003cs:*
11: Geneseqn2003ds:*
12: Geneseqn2004as:*
13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3300	100.0	3300	4 AAD11590	Aad11590 Partial s
2	3216.8	97.5	3300	4 AAD11592	Aad11592 Partial s
3	3207.2	97.2	3300	4 AAD11588	Aad11588 Partial s
4	3207.2	97.2	3300	4 AAD11589	Aad11589 Partial s
5	3207.2	97.2	3300	4 AAD11591	Aad11591 Partial s
6	2705.8	82.0	8101	4 AAD25681	Aad25681 Nucleotid
7	2705.8	82.0	8101	6 ABL53238	ABL53238 Nucleotid
8	2690.4	81.5	7259	12 ADO23608	Ado23608 DNA encod
9	2690.4	81.5	7322	12 ADO23613	Ado23613 DNA encod
10	2690.4	81.5	7352	12 ADO23603	Ado23603 DNA encod
11	2690.4	81.5	7370	12 ADO23598	Ado23598 DNA encod
12	2690.4	81.5	7370	12 ADO23639	Ado23639 DNA encod
13	2690.4	81.5	7370	12 ADO23649	Ado23649 DNA encod
14	2690.4	81.5	7370	12 ADO23644	Ado23644 DNA encod
15	2690.4	81.5	7403	12 ADO23594	Ado23594 DNA encod
16	2690.4	81.5	7442	12 ADO23590	Ado23590 DNA encod
17	2690.4	81.5	7478	12 ADO23584	Ado23584 DNA encod
18	2688.8	81.5	7112	12 ADO23618	Ado23618 DNA encod
19	2636.8	79.9	6806	6 ABL49925	ABL49925 Maltose b
20	2605.4	79.0	7553	12 ADO23588	Ado23588 DNA encod

21	1609.4	48.8	5558	12 ADL90419	Adl90419 Clostridi
22	1474	44.7	4700	9 ACF06053	Act06053 Vector pJ
23	1472.4	44.6	4700	12 ADL18670	Adl18670 Vector pJ
24	1462.4	44.3	4920	10 ADH73599	Adh73599 Plasmid p
25	1462.4	44.3	4935	10 ADH73598	Adh73598 Plasmid p
26	1462.4	44.3	4945	10 ADH73600	Adh73600 Plasmid p
27	1462.4	44.3	4951	10 ADH73601	Adh73601 Plasmid p
28	1418.2	43.0	5903	8 ABZ23939	Abz23939 Nucleotid
29	1417.6	43.0	1511	6 ABQ73200	Abq73200 E. coli t
c 30	1417.6	43.0	5926	2 AAV32977	Aav32977 Tn7 donor
c 31	1417.6	43.0	5926	6 AAD45059	Aad45059 Transposo
c 32	1417.6	43.0	5926	12 ADG46817	Adg46817 Donor pla
33	1362	41.3	5201	12 ADL72229	Adl72229 DNA seque
34	1362	41.3	5201	12 ADL72228	Adl72228 DNA seque
c 35	1238.4	37.5	5024	9 ACF06051	Act06051 Plasmid p
c 36	1238.4	37.5	5024	12 ADL18668	Adl18668 Plasmid p
37	1237.2	37.5	1922	6 ABQ73089	Abq73089 MBP and z
38	1237.2	37.5	1922	10 ADF83622	Adf83622 MBP-human
39	1204	36.5	4557	2 AAT90491	Aat90491 Vector pl
40	1204	36.5	4969	12 ADO33522	Ado33522 Expressio
41	1204	36.5	5006	10 ADE24112	Ade24112 plasmid p
42	1204	36.5	5064	12 ADO33532	Ado33532 Vector pG
43	1204	36.5	5082	12 ADO33531	Ado33531 Vector pG
44	1204	36.5	6259	6 AAD29720	Aad29720 Plasmid p
45	1204	36.5	6823	3 AAC55459	Aac55459 Destinati

ALIGNMENTS

RESULT 1
AAD11590
ID AAD11590 standard; DNA; 3300 BP.
XX AAD11590;
XX
DT 24-SEP-2001 (first entry)
XX
DE Partial sequence of pMal-m3 and ZFPm3 DNA.
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
KW modulation; plant technology; agriculture; ds.
XX
OS Unidentified.
XX
FH Key
CDS Location/Qualifiers
FT
FT /tag= a
FT /product= "ZFPm3 protein"
FT /note= "CDS does not include start and stop codon"
FT /partial
FT primer_bind 2740..2790
FT /tag= b
FT /bound_moiety= "F1-f2 primer"
FT primer_bind 2770..2850
FT /tag= c
FT /bound_moiety= "F1-f1 primer"
FT primer_bind 2824..2889
FT /tag= d
FT /bound_moiety= "F2-b primer"
FT primer_bind 2867..2940
FT /tag= e
FT /bound_moiety= "F2-f primer"
FT primer_bind 2916..2973
FT /tag= f
FT /bound_moiety= "F3-b1 primer"
FT primer_bind 2953..3021
FT /tag= g
FT /bound_moiety= "F3-b2 primer"
FT primer_bind 2992..3042
FT /tag= h
FT /bound_moiety= "F4-f2 primer"
FT primer_bind 3022..3102

```

FT      /*tag= i
FT      /bound_moiety= "F4-f1 primer"
FT      primer_bind      3076..3141
FT      /*tag= j
FT      /bound_moiety= "F5-b primer"
FT      primer_bind      3119..3192
FT      /*tag= k
FT      /bound_moiety= "F5-f primer"
FT      primer_bind      3168..3225
FT      /*tag= l
FT      /bound_moiety= "F6-b1 primer"
FT      primer_bind      3205..3273
FT      /*tag= m
FT      /bound_moiety= "F6-b2 primer"
XX
PN      WO200152620-A2.
PD      26-JUL-2001.
XX
PF      19-JAN-2001; 2001WO-US001817.
XX
PR      21-JAN-2000; 2000US-0177468P.
PR      21-JUL-2000; 2000US-00620897.
XX
PA      (SCRI ) SCRIPPS RES INST.
PA      (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
PI      Barbas CF, Stege JT, Guan X, Dalmia B;
XX      WPI; 2001-465325/50.
DR      P-PSDB; AAE06003.
XX
XX      New zinc finger proteins, useful for modulating or regulating gene
PT      expression and metabolic pathways in plants, e.g. for treating in the
PT      plant cells a disorder that is associated with abnormal expression of the
PT      target gene.
XX
PS      Example 4; Page 143-145; 156pp; English.
XX
CC      The patent discloses methods and compositions to modulate the expression
CC      of a target gene in plant cells. The method involves providing plant
CC      cells with a zinc finger protein (ZFP) which is capable of specifically
CC      binding to a target nucleotide sequence or its complementary strand
CC      within a target gene and allowing the ZFP binding to the target
CC      nucleotide sequence. Where the expression of the target gene in the plant
CC      cells is modulated. The ZFP and fusions of the ZFP proteins are useful
CC      for modulating or regulating gene expression and metabolic pathways in
CC      plants. The ZFP, fusion proteins and methods are useful in plant and
CC      agricultural technology. The method is useful particularly for treating a
CC      disorder in the plant cells, where the disorder is associated with
CC      abnormal expression of the target gene. The present DNA sequence is the
CC      partial sequence of pMal-m3 and ZFPm3 DNA
XX
SQ      Sequence 3300 BP; 845 A; 878 C; 877 G; 700 T; 0 U; 0 Other;
Query Match      100.0%; Score 3300; DB 4; Length 3300;
Best Local Similarity      100.0%; Pred. No. 0;
Matches 3300; Conservative      0; Mismatches      0; Indels      0; Gaps      0;
QY      1      CCGACACCATCGAATGGTGCAGAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA      60
DB      1      CCGACACCATCGAATGGTGCAGAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA      60
QY      61      GTCAATTCAGGGTGGTGAATGTGAACAGTAACGTTATACGATGTGCGAGAGTATGCCG      120
DB      61      GTCAATTCAGGGTGGTGAATGTGAACAGTAACGTTATACGATGTGCGAGAGTATGCCG      120
QY      121      GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCAGTCTTCGCGAAA      180
DB      121      GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCAGTCTTCGCGAAA      180
QY      181      CGCGGAAAAAGTGAAGCGCGCATGCGGAGCTGAATTACATTCACCAACCGCGTGGCAC      240
DB      181      CGCGGAAAAAGTGAAGCGCGCATGCGGAGCTGAATTACATTCACCAACCGCGTGGCAC      240

```

```

DB      181      CGCGGAAAAAGTGAAGCGCGCATGCGGAGCTGAATTACATTCACCAACCGCGTGGCAC      240
QY      241      AACAACTGGCGGGCAAAACAGTTCGTTGCTGATTTGGCGTTGCACCTCAGTCTGGCCCTGCG      300
DB      241      AACAACTGGCGGGCAAAACAGTTCGTTGCTGATTTGGCGTTGCACCTCAGTCTGGCCCTGCG      300
QY      301      ACGCGCGTTCGCAAAATTTGTGCGCGGATTAATCTCAGCGCGGATCAACTGGTTCGCCAGCG      360
DB      301      ACGCGCGTTCGCAAAATTTGTGCGCGGATTAATCTCAGCGCGGATCAACTGGTTCGCCAGCG      360
QY      361      TGGTGGTTCGATCGTAGAACGAAAGCGCGTTCGAAGCTCTGTAAAGCGCGGTGCACAATC      420
DB      361      TGGTGGTTCGATCGTAGAACGAAAGCGCGTTCGAAGCTCTGTAAAGCGCGGTGCACAATC      420
QY      421      TTCTCGCGCAACCGGTTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACCAAGATGCCA      480
DB      421      TTCTCGCGCAACCGGTTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACCAAGATGCCA      480
QY      481      TTGCTGTGGAAGCTGCCTGCACATAATCTTCGGGGTTATTTCTTGATGTCTCTGACCAGA      540
DB      481      TTGCTGTGGAAGCTGCCTGCACATAATCTTCGGGGTTATTTCTTGATGTCTCTGACCAGA      540
QY      541      CACCATCAACAGTAATTTTCTCCATGAAGACGGTACGCGACTGGCGTGGAGCATC      600
DB      541      CACCATCAACAGTAATTTTCTCCATGAAGACGGTACGCGACTGGCGTGGAGCATC      600
QY      601      TGGTCGCAATGGGTTCACCGCAAAATCGCGCTGTAGCGGGCCCATTAAGTCTGTCTCGG      660
DB      601      TGGTCGCAATGGGTTCACCGCAAAATCGCGCTGTAGCGGGCCCATTAAGTCTGTCTCGG      660
QY      661      CGGTCCTGCGTCTGGCTGGCGTGGCAATAATCTCACTCGCAATCAAAATCAGCCGATAG      720
DB      661      CGGTCCTGCGTCTGGCTGGCGTGGCAATAATCTCACTCGCAATCAAAATCAGCCGATAG      720
QY      721      CGGAACGGGAAGCGGACTGAGTGCCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA      780
DB      721      CGGAACGGGAAGCGGACTGAGTGCCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA      780
QY      781      ATGAGGCGATCTGTTCCCACTCGCATGCTGTTGCCAACGATCAGATGGCGCTGGCGCAA      840
DB      781      ATGAGGCGATCTGTTCCCACTCGCATGCTGTTGCCAACGATCAGATGGCGCTGGCGCAA      840
QY      841      TGGCGGCCAATTAACGAGTCCGGCTCGCGTTCGGTTCGGATATCTCGTAGTGGGATACG      900
DB      841      TGGCGGCCAATTAACGAGTCCGGCTCGCGTTCGGTTCGGATATCTCGTAGTGGGATACG      900
QY      901      ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAACCAACCATCAACAGGATTTTC      960
DB      901      ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAACCAACCATCAACAGGATTTTC      960
QY      961      GCCTGCTGGGGCAAAACAGCGTTCGACCGCTTGTGCAACTCTCTCAGGGCGAGCGGTGA      1020
DB      961      GCCTGCTGGGGCAAAACAGCGTTCGACCGCTTGTGCAACTCTCTCAGGGCGAGCGGTGA      1020
QY      1021      AGGGCAATCAGCTGTTGGCCGCTCTCACTGGTGAAGAAAAAACCCACCTCTGGCGCCCAATA      1080
DB      1021      AGGGCAATCAGCTGTTGGCCGCTCTCACTGGTGAAGAAAAAACCCACCTCTGGCGCCCAATA      1080
QY      1081      CGCAAAACCCCTCTCCCGCGGTTGGCGGATTCATTAAATGAGCTGGCAGCGACGAGTTT      1140
DB      1081      CGCAAAACCCCTCTCCCGCGGTTGGCGGATTCATTAAATGAGCTGGCAGCGACGAGTTT      1140
QY      1141      CCGCATGGAAGCGGGCAGTGAAGCGCAACGCAATTAATGTGAGTTAGCTCACTCAATTAG      1200
DB      1141      CCGCATGGAAGCGGGCAGTGAAGCGCAACGCAATTAATGTGAGTTAGCTCACTCAATTAG      1200
QY      1201      GCAGAACTTCATGTTTGCAGCTTATCATGCACTGCAGCGGTGCACCAATGCTTCTGGCG      1260
DB      1201      GCAGAACTTCATGTTTGCAGCTTATCATGCACTGCAGCGGTGCACCAATGCTTCTGGCG      1260
QY      1261      TCAGGCAAGCATCGGAAGCTGTGGTATGGCTGTGCAAGGTGCTGAAATCACTGGCAATATTCG      1320
DB      1261      TCAGGCAAGCATCGGAAGCTGTGGTATGGCTGTGCAAGGTGCTGAAATCACTGGCAATATTCG      1320

```

QY	1321	TGTCGCTCAAGCGCGACCTCCCGTTCGATATATGTTTTTGGCCGACATCATACGGTT	1380
DB	1321	TGTCGCTCAAGCGCGACCTCCCGTTCGATATATGTTTTTGGCCGACATCATACGGTT	1380
QY	1381	CTGGCAATATTTCTCAAAATGAGCTGTTGACAAATTAATCATCGCTCGTATATGTTGGA	1440
DB	1381	CTGGCAATATTTCTCAAAATGAGCTGTTGACAAATTAATCATCGCTCGTATATGTTGGA	1440
QY	1441	ATTGTGAGCGGATAACAATTTTACACAGAGAAACAGCCAGTCGGTTAGGTGTTTTTCAG	1500
DB	1441	ATTGTGAGCGGATAACAATTTTACACAGAGAAACAGCCAGTCGGTTAGGTGTTTTTCAG	1500
QY	1501	GCATTTCAACCAACAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAATCTGG	1560
DB	1501	GCATTTCAACCAACAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAATCTGG	1560
QY	1561	ATTAACGCGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
DB	1561	ATTAACGCGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTTCCACAGGTT	1680
DB	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTTCCACAGGTT	1680
QY	1681	GGGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACGACCGCTTTGGTGCTAC	1740
DB	1681	GGGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACGACCGCTTTGGTGCTAC	1740
QY	1741	GCTCAATCTGGGCTGTGTGCTGAATCACTCCCGGACAAAGCGTTCAGACAAAGCTGAT	1800
DB	1741	GCTCAATCTGGGCTGTGTGCTGAATCACTCCCGGACAAAGCGTTCAGACAAAGCTGAT	1800
QY	1801	CCGTTTACCTGGGATGCGGTACGTTTACAAAGCAAGCTGATGCTTTACCCGATCGCTGT	1860
DB	1801	CCGTTTACCTGGGATGCGGTACGTTTACAAAGCAAGCTGATGCTTTACCCGATCGCTGT	1860
QY	1861	GAAGCGTTATCGCTGATTTAACAAGATCTGCTGCGGACCCGCCAATAACCTGGGAA	1920
DB	1861	GAAGCGTTATCGCTGATTTAACAAGATCTGCTGCGGACCCGCCAATAACCTGGGAA	1920
QY	1921	GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAGGTGAAGCGCGCTGATGTTCAAC	1980
DB	1921	GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAGGTGAAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGCTACTTCACTGGCGCTGATGCTGCTGACGGGGTTATCGCTTCAAG	2040
DB	1981	CTGCAAGAACCGCTACTTCACTGGCGCTGATGCTGCTGACGGGGTTATCGCTTCAAG	2040
QY	2041	TATGAAAAACGCAAGTACGACATTAAGACGTTGGCGTGGATTAACGCTGGCGCAAGCG	2100
DB	2041	TATGAAAAACGCAAGTACGACATTAAGACGTTGGCGTGGATTAACGCTGGCGCAAGCG	2100
QY	2101	GCTCTGACCTTCTGGTTGACCTGATTAACCAACACATCAATGCAGACACCGATTAC	2160
DB	2101	GCTCTGACCTTCTGGTTGACCTGATTAACCAACACATCAATGCAGACACCGATTAC	2160
QY	2161	TCCATCGCAAGAGCTGCCCTTTAATAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG	2220
DB	2161	TCCATCGCAAGAGCTGCCCTTTAATAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG	2220
QY	2221	GCATGGTCAACATCCAAACCGTTGTTGGCTGCTGAGCGCAGGTATTAAACCGCCAGTTC	2280
DB	2221	GCATGGTCAACATCCAAACCGTTGTTGGCTGCTGAGCGCAGGTATTAAACCGCCAGTTC	2280
QY	2281	AGGGTCAACATCCAAACCGTTGTTGGCTGCTGAGCGCAGGTATTAAACCGCCAGTTC	2340
DB	2281	AGGGTCAACATCCAAACCGTTGTTGGCTGCTGAGCGCAGGTATTAAACCGCCAGTTC	2340
QY	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACTATCTGCTACTGATGAAGTCTG	2400
DB	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACTATCTGCTACTGATGAAGTCTG	2400

QY	2401	GAAGCGTTAATAAGACAAACCGCTGGTGGCTAGCGCTGAAGTCTTACGAGGAAG	2460
DB	2401	GAAGCGTTAATAAGACAAACCGCTGGTGGCTAGCGCTGAAGTCTTACGAGGAAG	2460
QY	2461	TTGGCGAAAGATCCACGTTATTGGCGCCACCATGAAAAACGCCAGAAAAGTGAATCATG	2520
DB	2461	TTGGCGAAAGATCCACGTTATTGGCGCCACCATGAAAAACGCCAGAAAAGTGAATCATG	2520
QY	2521	CCGAACATCCCGCAGATGTCGGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAACGCC	2580
DB	2521	CCGAACATCCCGCAGATGTCGGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAACGCC	2580
QY	2581	CCGAGCGTCTGACGATGTCGATGAAGCCCTGAAAAGCGCGACACTAAATTCGAGCTCG	2640
DB	2581	CCGAGCGTCTGACGATGTCGATGAAGCCCTGAAAAGCGCGACACTAAATTCGAGCTCG	2640
QY	2641	AACAACAAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTTCAAGATTC	2700
DB	2641	AACAACAAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTTCAAGATTC	2700
QY	2701	GGATCTCTTCTCTGTGGCCAGCGGCCCTTCGAGCCCGGGGAGAACCTTATGCTTGT	2760
DB	2701	GGATCTCTTCTCTGTGGCCAGCGGCCCTTCGAGCCCGGGGAGAACCTTATGCTTGT	2760
QY	2761	CGGNAATGTGTTAAGTCTCTCAGCGATCCTGGCCACCTGTTTCGCCACACGCGTACCAC	2820
DB	2761	CGGNAATGTGTTAAGTCTCTCAGCGATCCTGGCCACCTGTTTCGCCACACGCGTACCAC	2820
QY	2821	ACGGTGAAAAACCGTATAAATGCGGATGCGGCAAACTCTTTAGCACACGCGGCTCC	2880
DB	2821	ACGGTGAAAAACCGTATAAATGCGGATGCGGCAAACTCTTTAGCACACGCGGCTCC	2880
QY	2881	CTGGTGCCCATCAACGCACTCATACTGCGGAGAACCATACAAATGTCCAGATGTGGC	2940
DB	2881	CTGGTGCCCATCAACGCACTCATACTGCGGAGAACCATACAAATGTCCAGATGTGGC	2940
QY	2941	AAGTCTTTTACGCCAGAGCTCCAGCTGCTGCGCCACCAACGTTACTCACACGGGAGAG	3000
DB	2941	AAGTCTTTTACGCCAGAGCTCCAGCTGCTGCGCCACCAACGTTACTCACACGGGAGAG	3000
QY	3001	CCCTATGCTGTGCGGAATGTGTAAGTCTTTCAGCCAGAGCAGCTCCTGCTGGGCCAC	3060
DB	3001	CCCTATGCTGTGCGGAATGTGTAAGTCTTTCAGCCAGAGCAGCTCCTGCTGGGCCAC	3060
QY	3061	CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTTAGT	3120
DB	3061	CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTTAGT	3120
QY	3121	GACTGCCGGACCTTGCTGCGCATCAACGCACTCATACTGCGGAGAACCATACAAATGT	3180
DB	3121	GACTGCCGGACCTTGCTGCGCATCAACGCACTCATACTGCGGAGAACCATACAAATGT	3180
QY	3181	CCAGATGTGGCAAGTCTTTCTCCCAATCCAGCCATCTGTCGGCACCAACGTTACTCAC	3240
DB	3181	CCAGATGTGGCAAGTCTTTCTCCCAATCCAGCCATCTGTCGGCACCAACGTTACTCAC	3240
QY	3241	ACCGGTAAAAAACTAGTGGCCAGCGCGCCAGTACCGCTACGAGTTCGCGACTACGCT	3300
DB	3241	ACCGGTAAAAAACTAGTGGCCAGCGCGCCAGTACCGCTACGAGTTCGCGACTACGCT	3300

RESULT 2

AAD11592
ID AAD11592 standard; DNA; 3300 BP.

XX AAD11592;

AC AAD11592;

XX 24-SEP-2001 (first entry)

DE Partial sequence of pMal-Ap3 and ZFPap3 DNA.

XX Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;

KW modulation; plant technology; agriculture; Ap3; APETALA3; ds.

```

XX OS Unidentified.
XX PH Key
XX FT CDS
XX FT Location/Qualifiers
XX FT 2719..3270
XX FT /tag= a
XX FT /product= "ZFPap3 protein"
XX FT /transl_except= (pos:3124..3126, aa:Ser)
XX FT /transl_except= (pos:3127..3129, aa:Ser)
XX FT /transl_except= (pos:3136..3138, aa:Val)
XX FT /note= "CDS does not include start and stop codon"
XX FT /partial
XX FT primer_bind
XX FT 2740..2790
XX FT /tag= b
XX FT /bound_moiety= "F1-f2 primer"
XX FT primer_bind
XX FT 2770..2850
XX FT /tag= c
XX FT /bound_moiety= "F1-f1 primer"
XX FT primer_bind
XX FT 2824..2889
XX FT /tag= d
XX FT /bound_moiety= "F2-b primer"
XX FT primer_bind
XX FT 2867..2940
XX FT /tag= e
XX FT /bound_moiety= "F2-f primer"
XX FT primer_bind
XX FT 2916..2973
XX FT /tag= f
XX FT /bound_moiety= "F3-b1 primer"
XX FT primer_bind
XX FT 2953..3021
XX FT /tag= g
XX FT /bound_moiety= "F3-b2 primer"
XX FT primer_bind
XX FT 2992..3042
XX FT /tag= h
XX FT /bound_moiety= "F4-f2 primer"
XX FT primer_bind
XX FT 3022..3102
XX FT /tag= i
XX FT /bound_moiety= "F4-f1 primer"
XX FT primer_bind
XX FT 3076..3141
XX FT /tag= j
XX FT /bound_moiety= "F5-b primer"
XX FT primer_bind
XX FT 3119..3192
XX FT /tag= k
XX FT /bound_moiety= "F5-f primer"
XX FT primer_bind
XX FT 3168..3225
XX FT /tag= l
XX FT /bound_moiety= "F6-b1 primer"
XX FT primer_bind
XX FT 3205..3273
XX FT /tag= m
XX FT /bound_moiety= "F6-b2 primer"
XX MO200152620-A2.
XX
XX PD 26-JUL-2001.
XX
XX PF 19-JAN-2001; 2001WO-US001817.
XX PF 21-JAN-2000; 2000US-0177468P.
XX PR 21-JUL-2000; 2000US-00620897.
XX
XX PA (SCRI ) SCRIPPS RES INST.
XX PA (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
XX PI Barbas CF, Stege JT, Guan X, Dalmia B;
XX
XX DR WPI; 2001-45325/50.
XX DR P-PSDB; AAE06005.
XX
XX PT New zinc finger proteins, useful for modulating or regulating gene
XX PT expression and metabolic pathways in plants, e.g. for treating in the
XX PT plant cells a disorder that is associated with abnormal expression of the
XX PT target gene.
XX
XX PS Example 4; Page 148-149; 156pp; English.
XX

```

```

CC The patent discloses methods and compositions to modulate the expression
CC of a target gene in plant cells. The method involves providing plant
CC cells with a zinc finger protein (ZFP) which is capable of specifically
CC binding to a target nucleotide sequence or its complementary strand
CC within a target gene and allowing the ZFP binding to the target
CC nucleotide sequence, where the expression of the target gene in the plant
CC cells is modulated. The ZFP and fusions of the ZFP proteins are useful
CC for modulating or regulating gene expression and metabolic pathways in
CC plants. The ZFP, fusion proteins and methods are useful in plant and
CC agricultural technology. The method is useful particularly for treating a
CC disorder in the plant cells, where the disorder is associated with
CC abnormal expression of the target gene. The present DNA sequence is the
CC partial sequence of pMal-Ap3 (APETALA3) and ZFPap3 DNA
XX
XX SQ Sequence 3300 BP; 853 A; 872 C; 876 G; 699 T; 0 U; 0 Other;
XX
Query Match 97.5%; Score 3216.8; DB 4; Length 3300;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 3248; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTCGCGGTATGGCATATAGCGCCCGGAAGAGA 60
DB 1 CCGACACCATCGAATGGTGCAGAAACCTTTCGCGGTATGGCATATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTTCAGGGTGGTGAATGTCAAAACAGTATACGATGTGCGAGAGTATCCG 120
DB 61 GTCAATTTCAGGGTGGTGAATGTCAAAACAGTATACGATGTGCGAGAGTATCCG 120
QY 121 GTGTCTTTATCAGACCGTTTCCCGGTGTGAACACGAGCCAGCCAGCTTTCGCGAAAA 180
DB 121 GTGTCTTTATCAGACCGTTTCCCGGTGTGAACACGAGCCAGCCAGCTTTCGCGAAAA 180
QY 181 CGCGGGAAGTGAAGCGCGGATGCGGAGCTGAATATCATTCACACCGCGTGGCAC 240
DB 181 CGCGGGAAGTGAAGCGCGGATGCGGAGCTGAATATCATTCACACCGCGTGGCAC 240
QY 241 AACAACTGGCGGCAACAGTCCGTTCGTATTCGGCGTTGCCACTTCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGCAACAGTCCGTTCGTATTCGGCGTTGCCACTTCAGTCTGGCCCTGC 300
QY 301 ACSCGCGTGCAGAAATTCGCGCGGATTAATTCCTCGCGCGATCAACTGGGTGCAGCG 360
DB 301 ACSCGCGTGCAGAAATTCGCGCGGATTAATTCCTCGCGCGATCAACTGGGTGCAGCG 360
QY 361 TGTGTGTGTCGATGTAGAACGAGCGGTGTAAGCGCTGTAAGCGGCGGTGCACATC 420
DB 361 TGTGTGTGTCGATGTAGAACGAGCGGTGTAAGCGCTGTAAGCGGCGGTGCACATC 420
QY 421 TTCTCGCGCAACCGTTCAGTGGGCTGATCAATTAATTCCTCGCTGGATGACCGAGATGCCA 480
DB 421 TTCTCGCGCAACCGTTCAGTGGGCTGATCAATTAATTCCTCGCTGGATGACCGAGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCATTAATTCCTCGCGGCTTATTTCTTGATGTCTTGACAGAGA 540
DB 481 TTGCTGTGGAAGCTGCTGCATTAATTCCTCGCGGCTTATTTCTTGATGTCTTGACAGAGA 540
QY 541 CACCCATCAACAGTATTTTCTCCCATGAAGACGCTACGCGACTGGGCGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTTTCTCCCATGAAGACGCTACGCGACTGGGCGTGGAGCATC 600
QY 601 TGTTCGCAATGGGTACACGCAAAATCGCGCTGTTAGCGGGGCCCATTAAGTTCGTCTCGG 660
DB 601 TGTTCGCAATGGGTACACGCAAAATCGCGCTGTTAGCGGGGCCCATTAAGTTCGTCTCGG 660
QY 661 CGGCTCTGCTGCTGGCTGGCTGAATATCTCACTCGCAATCAAAATTCAGCCCATAG 720
DB 661 CGGCTCTGCTGCTGGCTGGCTGAATATCTCACTCGCAATCAAAATTCAGCCCATAG 720
QY 721 CGGAACGGGAAGCGGCTGAGTGCATGTCGGGTTTTTCAACAAACCATGCAAAATCTGA 780
DB 721 CGGAACGGGAAGCGGCTGAGTGCATGTCGGGTTTTTCAACAAACCATGCAAAATCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTGCGTGTGTCGCAACGATCAGATGGCGCTGGGCGCAA 840

```

||||| 781 ATGAGGGCATCGTCCCACTCGGATGCTGGTTGGTTGCCAACGATCAGATGGCGTGGGCGCAA 840
Qy TCGCGCCATTACCGAGTCCGGGCTCGCGGTGGTGGCGATATCTCGTAGTGGGATACG 900
Db TCGCGCCATTACCGAGTCCGGGCTCGCGGTGGTGGCGATATCTCGTAGTGGGATACG 900
Qy ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACACACATCAAAACAGGATTTTC 960
Db ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACACACATCAAAACAGGATTTTC 960
Qy GCCTGCTGGGCAACACAGCTGAGCCGCTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db GCCTGCTGGGCAACACAGCTGAGCCGCTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Qy AGGCAATCAGCTGTTGGCCGCTCTCACTGGTGAAGAAACACACCCCTGGCGCCCAATA 1080
Db AGGCAATCAGCTGTTGGCCGCTCTCACTGGTGAAGAAACACACCCCTGGCGCCCAATA 1080
Qy CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAGCACAGGTTT 1140
Db CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAGCACAGGTTT 1140
Qy CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTAG 1200
Db CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTAG 1200
Qy GCACAATTTCTATGTTTGACAGCTTATCATCGACTGCGGTCGCAACCAATGCTTCTGGCG 1260
Db GCACAATTTCTATGTTTGACAGCTTATCATCGACTGCGGTCGCAACCAATGCTTCTGGCG 1260
Qy TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAATCACTGATTAATTCG 1320
Db TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAATCACTGATTAATTCG 1320
Qy TGTGCTCAAGCGGCACCTCCGTTCTGGATATGTTTTTGGCGGACATCATACGGTT 1380
Db TGTGCTCAAGCGGCACCTCCGTTCTGGATATGTTTTTGGCGGACATCATACGGTT 1380
Qy CTGGCAATATTTCTGAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
Db CTGGCAATATTTCTGAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
Qy ATTGTAGCGGATAACAATTTACACAGAAACAGCCAGTCCGTTTAGGTGTTTTCAGA 1500
Db ATTGTAGCGGATAACAATTTACACAGAAACAGCCAGTCCGTTTAGGTGTTTTCAGA 1500
Qy GCACCTCACCAACAGGACCATAGATTATGAACCTGAAGAGGTAAACTGTTATCTGG 1560
Db GCACCTCACCAACAGGACCATAGATTATGAACCTGAAGAGGTAAACTGTTATCTGG 1560
Qy ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Db ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Qy ACCGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCGAAGAGAAATTCACAGGTT 1680
Db ACCGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCGAAGAGAAATTCACAGGTT 1680
Qy GCGGCAATGCGGATGGCCCTGACATTATCTTCTGGGCAACGACCGCTTGGTGGCTAC 1740
Db GCGGCAATGCGGATGGCCCTGACATTATCTTCTGGGCAACGACCGCTTGGTGGCTAC 1740
Qy GCTCAATCTGGCCTGTGTGGCTGAATCAACCGGCAAAAGCGTTCCAGSACAAGCTGTAT 1800
Db GCTCAATCTGGCCTGTGTGGCTGAATCAACCGGCAAAAGCGTTCCAGSACAAGCTGTAT 1800
Qy CGGTTTACCTGGGATGCGGTACGTTTACAAACGGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
Db CGGTTTACCTGGGATGCGGTACGTTTACAAACGGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
Qy GAAGCGTTATCGCTGATTTATAAAGAGATCTGCTGCCGAAACCGGCCAAAAACCTGGGAA 1920
Db GAAGCGTTATCGCTGATTTATAAAGAGATCTGCTGCCGAAACCGGCCAAAAACCTGGGAA 1920

Db GAAGCGTTATCGCTGATTTATAAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
Qy GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGTGAAGAGCGGCTGATGTTCAAC 1980
Db GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGTGAAGAGCGGCTGATGTTCAAC 1980
Qy CTGCAAGAACCGTACTTTCACCTGGCGCTGATTCCTGCTGACGGGGTTATGCGTTCAAG 2040
Db CTGCAAGAACCGTACTTTCACCTGGCGCTGATTCCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy TATCAAAAACGGCAAGTACGACATTAAGAACCTGGGCGTGGATAAACGCTGGCGCGAAACGG 2100
Db TATCAAAAACGGCAAGTACGACATTAAGAACCTGGGCGTGGATAAACGCTGGCGCGAAACGG 2100
Qy GGTCTGACCTTCTCGTGGTGAACCTGATTAATAAAACAAACACATGAATGCAGACACCGATTAC 2160
Db GGTCTGACCTTCTCGTGGTGAACCTGATTAATAAAACAAACACATGAATGCAGACACCGATTAC 2160
Qy TCCATCCGAGAGCTGCTTAAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db TCCATCCGAGAGCTGCTTAAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy GCATGGTCCAAATCGACACACAGCAAAAGTGAATTAATGGTGAACCGTACTGCGGACCTTC 2280
Db GCATGGTCCAAATCGACACACAGCAAAAGTGAATTAATGGTGAACCGTACTGCGGACCTTC 2280
Qy AAGGTCAACCATCAAAACCGTTCTGTTGGCGTGTGAGCGAGGTATTAACGCCGCCAGT 2340
Db AAGGTCAACCATCAAAACCGTTCTGTTGGCGTGTGAGCGAGGTATTAACGCCGCCAGT 2340
Qy CCGAACAAAGAGCTGGCAAAAAGTCTCTCGAAAACCTATCTGCTGACTGATGAAGTCTG 2400
Db CCGAACAAAGAGCTGGCAAAAAGTCTCTCGAAAACCTATCTGCTGACTGATGAAGTCTG 2400
Qy GAAGCGGTATAAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTTACGAGGAAG 2460
Db GAAGCGGTATAAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTTACGAGGAAG 2460
Qy TTGCGAAAGATCCAAGTATTTGCCGCCACCATGGAACAGCCCGAGAAAGTGAATCATG 2520
Db TTGCGAAAGATCCAAGTATTTGCCGCCACCATGGAACAGCCCGAGAAAGTGAATCATG 2520
Qy CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Qy GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Qy AACAAACAAACAAATTAACAAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db AACAAACAAACAAATTAACAAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Qy GGATCTCTTCTCTGTTGGCCAGCGGCCCTCGAGCCCGGGGAGAACCTTATCTGTT 2760
Db GGATCTCTTCTCTGTTGGCCAGCGGCCCTCGAGCCCGGGGAGAACCTTATCTGTT 2760
Qy CCGGAATGTGTAGTCTCTTCCAGCGATCTGCGCACCTGGTTTCGCGCACCGCGTACCAC 2820
Db CCGGAATGTGTAGTCTCTTCCAGCGAGACTCTTCTGTTGGCGCACCGAGGTACCAC 2820
Qy ACGGTGAAAAACCGTATAAATGCGCAGAGTGGCAAAATCTTTTAGCACAGCGGCTCC 2880
Db ACGGTGAAAAACCGTATAAATGCGCAGAGTGGCAAAATCTTTTAGCACAGGCTCC 2880
Qy CTGGTGGCCATCAACCGACTCATCTGCGAAGAACCATATAAATGTCCAGATGTGGC 2940
Db CTGGTGGCCATCAACCGACTCATCTGCGAAGAACCATATAAATGTCCAGATGTGGC 2940
Qy AAGTCTTTTTCAGCCAGCTCCAGCTGTTGGCGCACCAACGCTACTCACACGGGAGAG 3000
Db AAGTCTTTTTCAGCCAGCTCCAGCTGTTGGCGCACCAACGCTACTCACACGGGAGAG 3000

```

QY 3001 CCCTATGCTGTCCGGAATGGTAAAGTCCTTCAGCCAGAGCAGCTCCCTGGTGGCCAC 3060
Dd |||||
QY 3001 CCCTATGCTGTCCGGAATGGTAAAGTCCTTCAGCCAGAGCAGCTCCCTGGTGGCCAC 3060
Dd |||||
QY 3061 CAGCGTACCCACACGGGTGAAAACCGTATAAATGCCAGAGTGGGCAAAATCTTTAGT 3120
Dd |||||
QY 3061 CAGCGTACCCACACGGGTGAAAACCGTATAAATGCCAGAGTGGGCAAAATCTTTAGC 3120
Dd |||||
QY 3121 GACTGCCCGGACCTTGTCTCGGCATCAACGCACCTATCTACTGGCGAGAAGCCATACAAATGT 3180
Dd |||||
QY 3121 CAGCGGCCACCTTGGACGGCATCAACGCACCTATCTACTGGCGAGAAGCCATACAAATGT 3180
Dd |||||
QY 3181 CCAGAAATGGCAAGTCTTTCTCCAAATCCAGCCATCTGTCGGCACCAACGTACTCAC 3240
Dd |||||
QY 3181 CCAGAAATGGCAAGTCTTTCTCAACTTCAGGCAACTTGGTCCGTCCACCAACGTACTCAC 3240
Dd |||||
QY 3241 ACCGTTAAAAAATAGTGGCCAGCGCGCCAGTACCGGTACGAGTTCGGACTAGCT 3300
Dd |||||
QY 3241 ACCGTTAAAAAATAGTGGCCAGCGCGCCAGTACCGGTACGAGTTCGGACTAGCT 3300
Dd |||||

RESULT 3
AAD11588
ID AAD11588 standard; DNA; 3300 BP.
AC AAD11588;
XX
DT 24-SEP-2001 (first entry)
XX
DE Partial sequence of pMal-m1 and ZFPm1 DNA.
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
KW modulation; plant technology; agriculture; ds.
OS Unidentified.
XX
PH Key
FT Location/Qualifiers
FT 2719..3270
FT /*tag= a
FT /product= "ZFPm1 protein"
FT /transl_except= (pos:2785..2787, aa:Asp)
FT /transl_except= (pos:2788..2790, aa:Pro)
FT /transl_except= (pos:2791..2793, aa:Gly)
FT /transl_except= (pos:2874..2874, aa:Arg)
FT /transl_except= (pos:2875..2877, aa:Ala)
FT /transl_except= (pos:2878..2880, aa:His)
FT /transl_except= (pos:2884..2886, aa:Glu)
FT /transl_except= (pos:2953..2955, aa:Gln)
FT /transl_except= (pos:2959..2961, aa:Ser)
FT /transl_except= (pos:3121..3123, aa:Arg)
FT /transl_except= (pos:3124..3126, aa:Ser)
FT /transl_except= (pos:3127..3129, aa:Asp)
FT /transl_except= (pos:3130..3132, aa:Asn)
FT /transl_except= (pos:3136..3138, aa:Val)
FT /transl_except= (pos:3139..3141, aa:Arg)
FT /transl_except= (pos:3205..3207, aa:Gln)
FT /transl_except= (pos:3208..3210, aa:Ala)
FT /transl_except= (pos:3211..3213, aa:Gly)
FT /transl_except= (pos:3214..3216, aa:His)
FT /transl_except= (pos:3220..3222, aa:Ala)
FT /transl_except= (pos:3223..3225, aa:Ser)
FT /note= "CDS does not include start and stop codon"
FT /partial
FT 2740..2790
FT /*tag= b
FT /bound_moiety= "F1-f2 primer"
FT 2770..2850
FT /*tag= c
FT /bound_moiety= "F1-f1 primer"
FT 2824..2889
FT /*tag= d
FT /bound_moiety= "F2-b primer"

```

```

FT primer_bind 2867..2940
FT /*tag= e
FT /bound_moiety= "F2-f primer"
FT 2916..2973
FT /*tag= f
FT /bound_moiety= "F3-b1 primer"
FT 2953..3021
FT /*tag= g
FT /bound_moiety= "F3-b2 primer"
FT 2992..3042
FT /*tag= h
FT /bound_moiety= "F4-f2 primer"
FT 3022..3102
FT /*tag= i
FT /bound_moiety= "F4-f1 primer"
FT 3076..3141
FT /*tag= j
FT /bound_moiety= "F5-b primer"
FT 3119..3192
FT /*tag= k
FT /bound_moiety= "F5-f primer"
FT 3188..3225
FT /*tag= l
FT /bound_moiety= "F6-b1 primer"
FT 3205..3273
FT /*tag= m
FT /bound_moiety= "F6-b2 primer"

WO200152620-A2.
XX
XX 26-JUL-2001.
XX
XX 19-JAN-2001; 2001WO-US001817.
XX
XX 21-JAN-2000; 2000US-0177468P.
XX 21-JUL-2000; 2000US-00620897.
XX
XX (SCRI ) SCRIPPS RES INST.
XX (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
XX Barbas CF, Stege JT, Guan X, Dalmia B;
XX WPI; 2001-465325/50.
XX P-PSDB; AAE06000.
XX
XX New zinc finger proteins, useful for modulating or regulating gene
XX expression and metabolic pathways in plants, e.g. for treating in the
XX plant cells a disorder that is associated with abnormal expression of the
XX target gene.
XX
XX Example 4; Page 138-139; 156pp; English.
XX
XX The patent discloses methods and compositions to modulate the expression
XX of a target gene in plant cells. The method involves providing plant
XX cells with a zinc finger protein (ZFP) which is capable of specifically
XX binding to a target nucleotide sequence or its complementary strand
XX within a target gene and allowing the ZFP binding to the target
XX nucleotide sequence, where the expression of the target gene in the plant
XX cells is modulated. The ZFP and fusions of the ZFP proteins are useful
XX for modulating or regulating gene expression and metabolic pathways in
XX plants. The ZFP, fusion proteins and methods are useful in plant and
XX agricultural technology. The method is useful particularly for treating a
XX disorder in the plant cells, where the disorder is associated with
XX abnormal expression of the target gene. The present DNA sequence is the
XX partial sequence of pMal-m1 and ZFPm1 DNA.
XX
XX Sequence 3300 BP; 847 A; 874 C; 876 G; 703 T; 0 U; 0 Other;
SQ

```

```

Query Match 97.2%; Score 3207.2; DB 4; Length 3300;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGCAAAACCTTCGCGGTATGCGATGATAGCCCGGAAGAGA 60

```


Db 1 CCGACCACTCGAATGGTGC AAAA CTTTCGCGGTATGGCATGATAGCGCCGGAAGAGA 60
Qy 61 GTCAATTCAGGGTGTGAATGTGAACACCAAGTAACCTTATACGATGTCGCAGAGTATGCCG 120
Db 61 GTCAATTCAGGGTGTGAATGTGAACACCAAGTAACCTTATACGATGTCGCAGAGTATGCCG 120
Qy 121 GTGTCTCTTATCAGACCGTTTCCTCCGGTGTGAACACCAAGCCAGCCAGCACTTTCCTGCGAAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCTCCGGTGTGAACACCAAGCCAGCCAGCACTTTCCTGCGAAAA 180
Qy 181 CGCGGAAAAAGTGAAGCGGATGGCGGAGCTGAATTAACATTCCTCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGGATGGCGGAGCTGAATTAACATTCCTCAACCGCGTGGCAC 240
Qy 241 AACCACTCGCGGCAAAACAGTCTGTGCTGATTTGGCGTTTGCCACCTCCAGTCTGGGCCCTGC 300
Db 241 AACCACTCGCGGCAAAACAGTCTGTGCTGATTTGGCGTTTGCCACCTCCAGTCTGGGCCCTGC 300
Qy 301 AGCGCCGTTCGCAAAATTTGTTCGCGCGATTAATCTTCGCGCGGATCAACTGGGTGCCAGCG 360
Db 301 AGCGCCGTTCGCAAAATTTGTTCGCGCGATTAATCTTCGCGCGGATCAACTGGGTGCCAGCG 360
Qy 361 TGGTGGTTCGATGCTAGATGGGCTGATCAATTAATCTTCGCGCGGATCAACTGGGTGCCAATC 420
Db 361 TGGTGGTTCGATGCTAGATGGGCTGATCAATTAATCTTCGCGCGGATCAACTGGGTGCCAATC 420
Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTTCGCGCGGATCAACTGGGTGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTTCGCGCGGATCAACTGGGTGCCA 480
Qy 481 TTGTGTGGAGTCCCTGCACTAATGTTTCGCGCGTATTTCTGATGCTCTGACCCAGA 540
Db 481 TTGTGTGGAGTCCCTGCACTAATGTTTCGCGCGTATTTCTGATGCTCTGACCCAGA 540
Qy 541 CACCCATCAACAGTATTTCTCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTCTCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTGCATTTGGTCCACGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCTGCTCGG 660
Db 601 TGGTGCATTTGGTCCACGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCTGCTCGG 660
Qy 661 CGCGTCTGCTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTCCTCAACAAACCATGCAATGCTGA 780
Db 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTCCTCAACAAACCATGCAATGCTGA 780
Qy 781 ATGAGGGCATCGTTCCTCACTGCGATGCTGTTGCCAACGATGCGCTGGGCGCAA 840
Db 781 ATGAGGGCATCGTTCCTCACTGCGATGCTGTTGCCAACGATGCGCTGGGCGCAA 840
Qy 841 TCGCGCCATTAACGAGTCCGGGCTCGCGTGTGGTGGGATATCTCGTGTGGGATAG 900
Db 841 TCGCGCCATTAACGAGTCCGGGCTCGCGTGTGGTGGGATATCTCGTGTGGGATAG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACACCATCAAAAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACACCATCAAAAGGATTTTC 960
Qy 961 GCCTGCTGGGCAACACGAGTGGACCGCTTCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGTGGACCGCTTCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGGCCGCTCACTGGTGAAGAAAAACCACTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCGCTCACTGGTGAAGAAAAACCACTGGCGGCCAATA 1080
Qy 1081 CGCAAAACGCTCTCTCCCGCGGCTTGGCGGATTCATTAATGCACTGGGCAACAGAGTTT 1140

Db 1081 CGCAAAACGCTCTCTCCCGCGGCTTGGCGGATTCATTAATGCACTGGCAACAGAGTTT 1140
Qy 1141 CCCGACTGGAAGGGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGGGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Qy 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAGGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAGGTGCACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGACCCATTCGGAAGCTGTGGTGTGAGGTGTCAGGTGCTAAATCACTGCATAAATTCG 1320
Db 1261 TCAGGACCCATTCGGAAGCTGTGGTGTGAGGTGTCAGGTGCTAAATCACTGCATAAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTCTTGATTAATGTTTTTTCGCGCAATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTTGATTAATGTTTTTTCGCGCAATCAATACGGTT 1380
Qy 1381 CTGGCAATATTTCTGAATGAGCTTTCACACAGGAAACAGCCAGTCCGTTTTCGCGCAATCAATACGGTT 1440
Db 1381 CTGGCAATATTTCTGAATGAGCTTTCACACAGGAAACAGCCAGTCCGTTTTCGCGCAATCAATACGGTT 1440
Qy 1441 ATTGTGAGCGGATAACAAATTTTCACACAGGAAACAGCCAGTCCGTTTTCGCGCAATCAATACGGTT 1500
Db 1441 ATTGTGAGCGGATAACAAATTTTCACACAGGAAACAGCCAGTCCGTTTTCGCGCAATCAATACGGTT 1500
Qy 1501 GCACCTTCAACCAAGGACCATAGATATATGAATACTGAAGAGGTAAATCTGCTAATCTCG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATATATGAATACTGAAGAGGTAAATCTGCTAATCTCG 1560
Qy 1561 ATTTAAGCGCGATAAAGGCTATAACCGTCTCGTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTTAAGCGCGATAAAGGCTATAACCGTCTCGTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTTAAAGTCAACCGTTGAGCATCCGATAACTGGAAGAGAAATTCCTCCAGGTT 1680
Db 1621 ACCGGAATTTAAAGTCAACCGTTGAGCATCCGATAACTGGAAGAGAAATTCCTCCAGGTT 1680
Qy 1681 CGGCAACTGCGATGGCCCTGACATATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGCGATGGCCCTGACATATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGCGCTGTTGGCTGAATTAACCCCGGCAAAAGGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTTGGCTGAATTAACCCCGGCAAAAGGTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACCTCGGATGCGGTACGTTACAAACGCGAGCTGATTGCTTACCGATCCCTGTT 1860
Db 1801 CCGTTTACCTCGGATGCGGTACGTTACAAACGCGAGCTGATTGCTTACCGATCCCTGTT 1860
Qy 1861 GAAGCGTTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCCGCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCCGCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGACGGGGTGTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGACGGGGTGTATGCGTTCAAG 2040
Qy 2041 TATGAAAAACGCAAGTACGATTAAGAACGCTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGATTAAGAACGCTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAATAACAAACACATGAATGCAAGACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAATAACAAACACATGAATGCAAGACCGATTAC 2160
Qy 2161 TCCATCGCAGAACGCTGCTCTTAAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAACGCTGCTCTTAAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220

```
QY 2221 GCATGGTCAACATCGACACCAAGAGTGAATTATGGTAAACGGTACTGCCGACCTTC 2280
Db |||||||
QY 2221 GCATGGTCAACATCGACACCAAGAGTGAATTATGGTAAACGGTACTGCCGACCTTC 2280
Db |||||||
QY 2281 AAGGCTCAACCATCAAAACCGTTCGTCGGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db |||||||
QY 2281 AAGGCTCAACCATCAAAACCGTTCGTCGGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db |||||||
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACGTGAAGGCTCTG 2400
Db |||||||
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACGTGAAGGCTCTG 2400
Db |||||||
QY 2401 GAAGCGGTAAATAAGACAAACCGCTGGTGCCTGAGCTGAAGTCTTACGAGGAAGAG 2460
Db |||||||
QY 2401 GAAGCGGTAAATAAGACAAACCGCTGGTGCCTGAGCTGAAGTCTTACGAGGAAGAG 2460
Db |||||||
QY 2461 TTGGCGAAGATCCACGATATTCGCCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db |||||||
QY 2461 TTGGCGAAGATCCACGATATTCGCCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db |||||||
QY 2521 CCGAACATCCGCGAGATGTCGCTTTCTGGTATGCGGTGCTACTCGCGTGATCAACGCC 2580
Db |||||||
QY 2521 CCGAACATCCGCGAGATGTCGCTTTCTGGTATGCGGTGCTACTCGCGTGATCAACGCC 2580
Db |||||||
QY 2581 GCCAGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2640
Db |||||||
QY 2581 GCCAGCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2640
Db |||||||
QY 2641 AACACACACACATTAACAATAACAACCTCCGGATCGAGGAGGAGGATTCAGAAATTC 2700
Db |||||||
QY 2641 AACACACACACATTAACAATAACAACCTCCGGATCGAGGAGGAGGATTCAGAAATTC 2700
Db |||||||
QY 2701 GGATCCTCTTCTGTGGCCAGCGGCCCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
Db |||||||
QY 2701 GGATCCTCTTCTGTGGCCAGCGGCCCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
Db |||||||
QY 2761 CCGGAATGTGTAAGTCTTCTGAGGATCCTGGCCACCTGTTCCGCCACCGAGCTACCCAC 2820
Db |||||||
QY 2761 CCGGAATGTGTAAGTCTTCTGAGGATCCTGGCCACCTGTTCCGCCACCGAGCTACCCAC 2820
Db |||||||
QY 2821 ACGGGTGAAAACCGTATAATGCCAGAGTGGCGGCAAACTTTTAGCACCAGCGGCTCC 2880
Db |||||||
QY 2821 ACGGGTGAAAACCGTATAATGCCAGAGTGGCGGCAAACTTTTAGCACCAGCTACCAAC 2880
Db |||||||
QY 2881 CTGTCGCGCCATCAACGACCTCATCTACTGGCGAGAGCCATACAATGTCAGAAATGTGGC 2940
Db |||||||
QY 2881 CTGTCGCGCCATCAACGACCTCATCTACTGGCGAGAGCCATACAATGTCAGAAATGTGGC 2940
Db |||||||
QY 2941 AAGTCTTTCTGCGGATGTGTAAGTCTTCTGAGCCAGAGCAGCTCCCTGGTGGCGCCAC 3000
Db |||||||
QY 2941 AAGTCTTTCTGCGGATGTGTAAGTCTTCTGAGCCAGAGCAGCTCCCTGGTGGCGCCAC 3000
Db |||||||
QY 3001 CCCTATGCTTGTCCGGATGTGTAAGTCTTCTGAGCCAGAGCAGCTCCCTGGTGGCGCCAC 3060
Db |||||||
QY 3001 CCCTATGCTTGTCCGGATGTGTAAGTCTTCTGAGCCAGAGCAGCTCCCTGGTGGCGCCAC 3060
Db |||||||
QY 3061 CAGCGTACCCACACCGGTGAAAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTAGT 3120
Db |||||||
QY 3061 CAGCGTACCCACACCGGTGAAAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTAGT 3120
Db |||||||
QY 3121 GACTGCGCGACCTTGTTCGCCATCAACGACCTCATCTGCGGAGAGGAGCCATACAATGT 3180
Db |||||||
QY 3121 GACTGCGCGACCTTGTTCGCCATCAACGACCTCATCTGCGGAGAGGAGCCATACAATGT 3180
Db |||||||
QY 3181 CCAGAAATGTGGCAAGTCTTTCTCCCAATCCAGCCATCTCGTCGGCACCAACGCTACTCAC 3240
Db |||||||
QY 3181 CCAGAAATGTGGCAAGTCTTTCTCCCAATCCAGCCATCTCGTCGGCACCAACGCTACTCAC 3240
Db |||||||
QY 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGGCGAGTACCCGTACGAGTTCGGGACTAGCT 3300
Db |||||||
QY 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGGCGAGTACCCGTACGAGTTCGGGACTAGCT 3300
Db |||||||
```

```
RESULT 4
RAD11589
ID AAD11589 standard; DNA; 3300 BP.
XX
AC AAD11589;
XX
DT 24-SEP-2001 (first entry)
XX
DE Partial sequence of pMal-m2 and ZFPm2 DNA.
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
modulation; plant technology; agriculture; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
CDS 2719..3270
FT /*tag= a
FT /product= "ZFPm2 protein"
FT /note= "CDS does not include start and stop codon"
FT /partial
FT primer_bind 2740..2790
FT /*tag= b
FT /bound_moiety= "F1-f2 primer"
FT primer_bind 2770..2850
FT /*tag= c
FT /bound_moiety= "F1-f1 primer"
FT primer_bind 2824..2889
FT /*tag= d
FT /bound_moiety= "F2-b primer"
FT primer_bind 2867..2940
FT /*tag= e
FT /bound_moiety= "F2-f primer"
FT primer_bind 2916..2973
FT /*tag= f
FT /bound_moiety= "F3-b1 primer"
FT primer_bind 2953..3021
FT /*tag= g
FT /bound_moiety= "F3-b2 primer"
FT primer_bind 2992..3042
FT /*tag= h
FT /bound_moiety= "F4-f2 primer"
FT primer_bind 3022..3102
FT /*tag= i
FT /bound_moiety= "F4-f1 primer"
FT primer_bind 3076..3141
FT /*tag= j
FT /bound_moiety= "F5-b primer"
FT primer_bind 3119..3192
FT /*tag= k
FT /bound_moiety= "F5-f primer"
FT primer_bind 3168..3225
FT /*tag= l
FT /bound_moiety= "F6-b1 primer"
FT primer_bind 3205..3273
FT /*tag= m
FT /bound_moiety= "F6-b2 primer"
XX
PN WO200152620-A2.
XX
PD 26-JUL-2001.
XX
XX 19-JAN-2001; 2001WO-US001817.
XX
XX 21-JAN-2000; 2000US-0177468P.
XX 21-JUL-2000; 2000US-00620897.
XX
(SCRI ) SCRIPPS RES INST.
PA (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
PI Barbas CF, Stege JT, Guan X, Dalmia B;
XX
```

DR WPI; 2001-465325/50.
DR P-PSDB; AAE06002.
XX
PT New zinc finger proteins, useful for modulating or regulating gene
PT expression and metabolic pathways in plants, e.g. for treating in the
PT plant cells a disorder that is associated with abnormal expression of the
PT target gene.
XX
PS Example 4; Page 140-142; 156pp; English.
XX
CC The patent discloses methods and compositions to modulate the expression
CC of a target gene in plant cells. The method involves providing plant
CC cells with a zinc finger protein (ZFP) which is capable of specifically
CC binding to a target nucleotide sequence or its complementary strand
CC within a target gene and allowing the ZFP binding to the target
CC nucleotide sequence, where the expression of the target gene in the plant
CC cells is modulated. The ZFP and fusions of the ZFP proteins are useful
CC for modulating or regulating gene expression and metabolic pathways in
CC plants. The ZFP, fusion proteins and methods are useful in plant and
CC agricultural technology. The method is useful particularly for treating a
CC disorder in the plant cells, where the disorder is associated with
CC abnormal expression of the target gene. The present DNA sequence is the
CC partial sequence of pMal-m2 and ZFPm2 DNA
XX
SQ Sequence 3300 BP; 847 A; 874 C; 876 G; 703 T; 0 U; 0 Other;

Query Match 97.2%; Score 3207.2; DB 4; Length 3300;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGATGGTGCAGAAACCTTTCGGCGTATGCGCATGATAGGCCCGGAAGAGA 60
Db 1 CCGACACCATCGATGGTGCAGAAACCTTTCGGCGTATGCGCATGATAGGCCCGGAAGAGA 60

Qy 61 GTCGAATTCAGGGTGTGAATGTGAACACAGTAACGTTATACGATCTCGCAGAGTATCGCG 120
Db 61 GTCGAATTCAGGGTGTGAATGTGAACACAGTAACGTTATACGATCTCGCAGAGTATCGCG 120

Qy 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACAGCCAGCCAGCCAGCTTTCGGAATA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACAGCCAGCCAGCCAGCTTTCGGAATA 180

Qy 181 CGCGGGAAAAAGTGAAGCGGGATGGCGAGCTGAATTAATCCCAACCGCGTGGCAC 240
Db 181 CGCGGGAAAAAGTGAAGCGGGATGGCGAGCTGAATTAATCCCAACCGCGTGGCAC 240

Qy 241 AACAACTGCGGGCAAAACAGTGTGCTGATTGGGGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGCGGGCAAAACAGTGTGCTGATTGGGGTTGGCCACCTCCAGTCTGGCCCTGC 300

Qy 301 ACGGCCGTTCGCAAAATTGTGCGGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
Db 301 ACGGCCGTTCGCAAAATTGTGCGGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCG 360

Qy 361 TGGTGGTGTGATGTAGAACGAAGCGGGTTCGAAGCGCTGTAAAGCGGGTGCACAATC 420
Db 361 TGGTGGTGTGATGTAGAACGAAGCGGGTTCGAAGCGCTGTAAAGCGGGTGCACAATC 420

Qy 421 TTCTCGCCCAACCGCTCAGTGGGCTGATTAATTAATCTCCGCTGATGACCAAGGATGCCA 480
Db 421 TTCTCGCCCAACCGCTCAGTGGGCTGATTAATTAATCTCCGCTGATGACCAAGGATGCCA 480

Qy 481 TTGCTGTGGAGCTGCCTGCACATAATGTTCGGCGTTATTTCTTGATGTCTCTGACCCAGA 540
Db 481 TTGCTGTGGAGCTGCCTGCACATAATGTTCGGCGTTATTTCTTGATGTCTCTGACCCAGA 540

Qy 541 CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACCGCATGCGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACCGCATGCGCGTGGAGCATC 600

Qy 601 TGGTGGCATTTGGTTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTCTGTCTCGG 660
Db 601 TGGTGGCATTTGGTTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTCTGTCTCGG 660

Qy 661 CGCGTCTGCGGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGCATAG 720
Db 661 CGCGTCTGCGGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGCATAG 720

Qy 721 CGGAACGGGAAGCGGACTGGAGTGCCATGTCGGGTTTTCACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGCGGACTGGAGTGCCATGTCGGGTTTTCACAAACCATGCAAAATGCTGA 780

Qy 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGGCAACGATCAGATGCGCTGGGCGCAA 840
Db 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGGCAACGATCAGATGCGCTGGGCGCAA 840

Qy 841 TGCGGGCATTTACCGAGTCCGGGCTGCGGTTGGTGCGGATATCTCGGTAGTGGGATAG 900
Db 841 TGCGGGCATTTACCGAGTCCGGGCTGCGGTTGGTGCGGATATCTCGGTAGTGGGATAG 900

Qy 901 ACGATACCGGAAGACAGCTCATGTTATATCCCGCGTTTAACCAACCATCAAAAGGATTTTC 960
Db 901 ACGATACCGGAAGACAGCTCATGTTATATCCCGCGTTTAACCAACCATCAAAAGGATTTTC 960

Qy 961 GCCTGCTGGGGCAAAACCAAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGTGA 1020
Db 961 GCCTGCTGGGGCAAAACCAAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGTGA 1020

Qy 1021 AGGGCAATCAGCTGTGTCCTCTCACTGTGTAAGAAAAACCAACCTTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGTCCTCTCACTGTGTAAGAAAAACCAACCTTGGCGCCCAATA 1080

Qy 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTAATTAATGAGCTGCGACGACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTAATTAATGAGCTGCGACGACAGGTTT 1140

Qy 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG 1200

Qy 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACGGTGCAACCAATGTTCTGCGG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACGGTGCAACCAATGTTCTGCGG 1260

Qy 1261 TCAGGCAGCCATCGGAAGCTGTGTTGCTGAGTAAATCACTGATCAATTAATTCG 1320
Db 1261 TCAGGCAGCCATCGGAAGCTGTGTTGCTGAGTAAATCACTGATCAATTAATTCG 1320

Qy 1321 TGTGCTCAAGCGCACCTCCCGTCTTGATTAATGTTTTTGGCCGACATATAACGGTT 1380
Db 1321 TGTGCTCAAGCGCACCTCCCGTCTTGATTAATGTTTTTGGCCGACATATAACGGTT 1380

Qy 1381 CTGGCAATATTTCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTATTAATGTGTGA 1440
Db 1381 CTGGCAATATTTCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTATTAATGTGTGA 1440

Qy 1441 ATTGTGAGCGGATTAACAAATTTTCAACAGGAACAGCCAGTCCGTTAGGTGTTTCACGA 1500
Db 1441 ATTGTGAGCGGATTAACAAATTTTCAACAGGAACAGCCAGTCCGTTAGGTGTTTCACGA 1500

Qy 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAGCTGAAGAGGTAATACTGTAATCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAGCTGAAGAGGTAATACTGTAATCTGG 1560

Qy 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTGGTAAAGAAATTCGAGAAAGAT 1620

Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCAGAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCAGAGGTT 1680

Qy 1681 GCGGCACTGCGGATGGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTGTGGCTAC 1740
Db 1681 GCGGCACTGCGGATGGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTGTGGCTAC 1740

```
QY 1741 GCTCAATCTGGCTGTGGCTGAATATCAACCCCGGACAAAGCGTTCACAGCAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAATATCAACCCCGGACAAAGCGTTCACAGCAAGCTGTAT 1800
QY 1801 CCGTTTACTCGGGATGCGGTAGTTTACACGCAAGCTGTATGCTTACCGCATCGCTGTT 1860
Db 1801 CCGTTTACTCGGGATGCGGTAGTTTACACGCAAGCTGTATGCTTACCGCATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCCCGCAAAACCTTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCCCGCAAAACCTTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAGAGAAAGTAAAGAGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAGAGAAAGTAAAGAGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTCTTCAACCTGGCGCTGATGCTGCTGACGGGGTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTCTTCAACCTGGCGCTGATGCTGCTGACGGGGTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGCTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGCTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGGTTGACCTGATTAATAAACHAAACATGAAATGACAGACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCGGTTGACCTGATTAATAAACHAAACATGAAATGACAGACCGGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAATGAAGCGCAACACAGCATGACCATCAAGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAATGAAGCGCAACACAGCATGACCATCAAGCGCCGCTGG 2220
QY 2221 GCATGGTCCAAATCGACACAGCAAAAGTGAATTAATGTTGTAACGGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACAGCAAAAGTGAATTAATGTTGTAACGGTACTGCCGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTGTTGGCTGTCTGAGCGCAGGTATTAACCGCGCCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTGTTGGCTGTCTGAGCGCAGGTATTAACCGCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAACGGTTTAATAAGACAAACCGCTGGTGCCTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAACGGTTTAATAAGACAAACCGCTGGTGCCTAGCGCTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCACGATTTGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTTGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTATGCGCTGCTGCTGCGGTGATCAAGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTATGCGCTGCTGCTGCGGTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGAGATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGAGATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAGGATTCAGAAATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAGGATTCAGAAATTC 2700
QY 2701 GGATCTCTTCTCTGTGGCCAGGCGGCTCGAGCCCGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTGGCCAGGCGGCTCGAGCCCGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGCTGAAGTCTTTCAGGATCTCTGGCCACCTGCTGCGCACAGCGTATCCAC 2820
Db 2761 CCGGAATGCTGAAGTCTTTCAGGATCTCTGCTGAGCTCTCACCTGGTGGCCACAGCGTATCCAC 2820
QY 2821 ACGGGTGAACAAACCGTATATAATCCAGAGTGGCGCAATCTTTTAGCACCGCGCTCC 2880
```

```
Db 2821 ACGGGTGAACAAACCGTATATAATGCCAGAGTGGCGAAATCTTTTAGCAGTCCAGCAAC 2880
QY 2881 CTGGTGGCCCATCAACGCACTCATATCTGGCGAGAGCCATACAAATGTCCAGAAATGTGGC 2940
Db 2881 CTGGTGGCCCATCAACGCACTCATATCTGGCGAGAGCCATACAAATGTCCAGAAATGTGGC 2940
QY 2941 AAGTCTTTTACGCGAGAGCTCCAGCGCTGGTGGCCACCAAGTACTCATACCGGGGAGAG 3000
Db 2941 AAGTCTTTTCTCGGTCTGCAATCTCTGTCGGGACCAACGTAATCTCATACCGGGGAGAG 3000
QY 3001 CCCTATCTTGTTCGGGAATGTGGTAACTCTTCTCAGCCAGAGCAGCTCCCTGGTGCGCCAC 3060
Db 3001 CCCTATCTTGTTCGGGAATGTGGTAACTCTTCTCAGCCAGAGCAGTAACTCTGGTGCGCCAC 3060
QY 3061 CAGGTTACCCACACGCGGTGAAAAACCTATAATATGCCCAGAGTGGCGCAATCTTTTAGT 3120
Db 3061 CAGGTTACCCACACGCGGTGAAAAACCTATAATATGCCCAGAGTGGCGCAATCTTTTAGT 3120
QY 3121 GACTGCCGCGACCTTGTCTGCCATCAACGCACTCATATCTGCGGACCAACCGTACTCTAC 3180
Db 3121 CAGCCGCGCCACCTGGCCAGCCATCAACGCACTCATATCTGCGGAGAGCCATACAAATGT 3180
QY 3181 CCAGAAATGTGGCAAGTCTTTTCTCCCAATCCAGCCATCTCTGCGGACCAACCGTACTCTAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTTTCTCTCGTCTGACAACTCTGTCGCGGACCAACCGTACTCTAC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCCAGGCGCGCAGTACCCGTACGAGTTCGCGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCCAGGCGCGCAGTACCCGTACGAGTTCGCGACTACGCT 3300

RESULT 5
AAD11591
ID AAD11591 standard; DNA; 3300 BP.
XX
AC AAD11591;
XX
DT 24-SBP-2001 (first entry)
XX
DE Partial sequence of pMal-m4 and ZFPm4 DNA.
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
XX modulation; plant technology; agriculture; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 2719..3270
FT /tag= a
FT /product= "ZFPm4 protein"
FT /trnsl_except= (pos:3046..3048, aa:Ser)
FT /note= "CDS does not include start and stop codon"
FT /partial
FT primer_bind 2740..2790
FT /tag= b
FT /bound_moiety= "F1-f2 primer"
FT primer_bind 2770..2850
FT /tag= c
FT /bound_moiety= "F1-f1 primer"
FT primer_bind 2824..2889
FT /tag= d
FT /bound_moiety= "F2-b primer"
FT primer_bind 2867..2940
FT /tag= e
FT /bound_moiety= "F2-f primer"
FT primer_bind 2916..2973
FT /tag= f
FT /bound_moiety= "F3-b1 primer"
FT primer_bind 2953..3021
FT /tag= g
FT /bound_moiety= "F3-b2 primer"
FT primer_bind 2992..3042
```

```
FT      /*tag= h
FT      /bound_moiety= "F4-f2 primer"
FT      3022..3102
FT      /*tag= i
FT      /bound_moiety= "F4-f1 primer"
FT      3076..3141
FT      /*tag= j
FT      /bound_moiety= "F5-b primer"
FT      3119..3192
FT      /*tag= k
FT      /bound_moiety= "F5-f primer"
FT      3168..3225
FT      /*tag= l
FT      /bound_moiety= "F6-b1 primer"
FT      3205..3273
FT      /*tag= m
FT      /bound_moiety= "F6-b2 primer"
XX
XX      WO200152620-A2.
XX
XX      26-JUL-2001.
XX
XX      19-JAN-2001; 2001WO-US001817.
XX
XX      21-JAN-2000; 2000US-0177468P.
XX      21-JUL-2000; 2000US-00620897.
XX
XX      (SCRI ) SCRIPPS RES INST.
XX      (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
XX      Barbás CF, Stege JT, Guan X, Dalmia B;
XX      WPI; 2001-465325/50.
XX      P-PSDB; AAE06004.
XX
XX      New zinc finger proteins, useful for modulating or regulating gene
XX      expression and metabolic pathways in plants, e.g. for treating in the
XX      plant cells a disorder that is associated with abnormal expression of the
XX      target gene.
XX
XX      Example 4; Page 145-147; 156pp; English.
XX
XX      The patent discloses methods and compositions to modulate the expression
XX      of a target gene in plant cells. The method involves providing plant
XX      cells with a zinc finger protein (ZFP) which is capable of specifically
XX      binding to a target nucleotide sequence or its complementary strand
XX      within a target gene and allowing the ZFP binding to the target
XX      nucleotide sequence, where the expression of the target gene in the plant
XX      cells is modulated. The ZFP and fusions of the ZFP proteins are useful
XX      for modulating or regulating gene expression and metabolic pathways in
XX      plants. The ZFP, fusion proteins and methods are useful in plant and
XX      agricultural technology. The method is useful particularly for treating a
XX      disorder in the plant cells, where the disorder is associated with
XX      abnormal expression of the target gene. The present DNA sequence is the
XX      partial sequence of pMal-m4 and ZFPm4 DNA
XX
XX      Sequence 3300 BP; 849 A; 865 C; 880 G; 706 T; 0 U; 0 Other;
XX
XX      Query Match      97.2%; Score 3207.2; DB 4; Length 3300;
XX      Best Local Similarity 98.2%; Pred. No. 0;
XX      Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
XX
XX      1  CCGACACCATCAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCCCGGAAGAGA 60
XX      |
XX      |
XX      1  CCGACACCATCAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCCCGGAAGAGA 60
XX      |
XX      |
XX      61  GTCAATTACGGGTGGTGAATGTGAACACCAAGTAAACGTTATACGATTCGACAGTATGCCG 120
XX      |
XX      |
XX      61  GTCAATTACGGGTGGTGAATGTGAACACCAAGTAAACGTTATACGATTCGACAGTATGCCG 120
XX      |
XX      |
XX      121  GTGTCCTTATCAGACCGGTTCCCGGGTGGTGAACACGAGCCAGCCAGTTTCTGCGAAAA 180
XX      |
XX      |
XX      121  GTGTCCTTATCAGACCGGTTCCCGGGTGGTGAACACGAGCCAGCCAGTTTCTGCGAAAA 180
XX      |
XX      |
```

```
QY      181  CGCGGAAAAAGTGGAAAGCGCGGATGGCGGAGCTGAATTATCATTTCCAAACCGGTGGCAC 240
DB      |
DB      181  CGCGGAAAAAGTGGAAAGCGCGGATGGCGGAGCTGAATTATCATTTCCAAACCGGTGGCAC 240
QY      241  AACAACTGGCGGCAAAACAGTCGTTGCTGATTGGCTTGCACCTCCAGTCTGGCCCTGC 300
DB      |
DB      241  AACAACTGGCGGCAAAACAGTCGTTGCTGATTGGCTTGCACCTCCAGTCTGGCCCTGC 300
QY      301  ACGCGCGTCGCAAAATTTGTTCGCGCGGATTAATCTTCGCGCGGATCAATCGGGTGCACGC 360
DB      |
DB      301  ACGCGCGTCGCAAAATTTGTTCGCGCGGATTAATCTTCGCGCGGATCAATCGGGTGCACGC 360
QY      361  TGGTGGTTCGATCGGTAGAACGAAGCGCGTCGAAGCTGTAAAGCGCGGTGCACAATC 420
DB      |
DB      361  TGGTGGTTCGATCGGTAGAACGAAGCGCGTCGAAGCTGTAAAGCGCGGTGCACAATC 420
QY      421  TTCTCGCGCAACCGCTCAGTGGGCTGATCACTAACTATCCGCTGGATCCGCTGATCACCAGGATGCCA 480
DB      |
DB      421  TTCTCGCGCAACCGCTCAGTGGGCTGATCACTAACTATCCGCTGGATCCGCTGATCACCAGGATGCCA 480
QY      481  TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
DB      |
DB      481  TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
QY      541  CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
DB      |
DB      541  CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
QY      601  TGGTGGCATTTGGGTCAACGCAAAATCGCGCTTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
DB      |
DB      601  TGGTGGCATTTGGGTCAACGCAAAATCGCGCTTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
QY      661  CGCGTCTGCGCTGCGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB      |
DB      661  CGCGTCTGCGCTGCGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY      721  CGGAAACGGGAAAGGCGACTGGAGTGCCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
DB      |
DB      721  CGGAAACGGGAAAGGCGACTGGAGTGCCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
QY      781  ATGAGGGCATGTTTCCCACTGGGATGCTGGTTCGCAACGATCAGATGCGGTGGCGGCA 840
DB      |
DB      781  ATGAGGGCATGTTTCCCACTGGGATGCTGGTTCGCAACGATCAGATGCGGTGGCGGCA 840
QY      841  TGGCGCCATTACCGAGTCCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
DB      |
DB      841  TGGCGCCATTACCGAGTCCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
QY      901  ACGATACCGAAGACAGCTCATGTTTATATPCCGCGGTTAAACCAACATCAAAACAGGATTTTC 960
DB      |
DB      901  ACGATACCGAAGACAGCTCATGTTTATATPCCGCGGTTAAACCAACATCAAAACAGGATTTTC 960
QY      961  GCCTGTGGGGCAACAGCGTGACCGCTTGTGCGAACTCTCTCAGGCGCCAGCGGTGA 1020
DB      |
DB      961  GCCTGTGGGGCAACAGCGTGAGCCGCTTGTGCGAACTCTCTCAGGCGCCAGCGGTGA 1020
QY      1021  AGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAACCAACCTTGGCGCCCAATA 1080
DB      |
DB      1021  AGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAACCAACCTTGGCGCCCAATA 1080
QY      1081  CGCAAAACCGCTCTCTCCCGCGGTTGGCCGATTCATTAAATGACGTGGCACACAGGTTT 1140
DB      |
DB      1081  CGCAAAACCGCTCTCTCCCGCGGTTGGCCGATTCATTAAATGACGTGGCACACAGGTTT 1140
QY      1141  CCAGCTGGAAGCGGCGAGTGAGGCAACGCAATTAATGTAGTTAGTCTACTCATTTAG 1200
DB      |
DB      1141  CCAGCTGGAAGCGGCGAGTGAGGCAACGCAATTAATGTAGTTAGTCTACTCATTTAG 1200
QY      1201  GCACAATTTCTATGTTTGACAGCTTATCATCGACTGACGGTGCACCAATGCTTCTGGCG 1260
DB      |
DB      1201  GCACAATTTCTATGTTTGACAGCTTATCATCGACTGACGGTGCACCAATGCTTCTGGCG 1260
```

QY 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTGAATCACTGCATAAATTCG 1320
Db 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTGAATCACTGCATAAATTCG 1320
QY 1321 TGTCCGCTCAAGCGGCACCTCCGCTTCGATATATGTTTTTGGCGCAGCATCAATACGGTT 1380
Db 1321 TGTCCGCTCAAGCGGCACCTCCGCTTCGATATATGTTTTTGGCGCAGCATCAATACGGTT 1380
QY 1381 CTGGCAAAATATTTCTGAATAGCTGTTGACATTAATCATCGGCTGTAATGTTGGA 1440
Db 1381 CTGGCAAAATATTTCTGAATAGCTGTTGACATTAATCATCGGCTGTAATGTTGGA 1440
QY 1441 ATTGTAGCGGATAACAATTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTCACGA 1500
Db 1441 ATTGTAGCGGATAACAATTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTCACGA 1500
QY 1501 GCACCTTCAACACAGGACCATAGATTATGAAAACCTGAAGAGGTAAACTCGGTAATCTGG 1560
Db 1501 GCACCTTCAACACAGGACCATAGATTATGAAAACCTGAAGAGGTAAACTCGGTAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGTTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGTTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAACTGGAAGAGAAATTCACACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAACTGGAAGAGAAATTCACACAGGTT 1680
QY 1681 CGGCACTGCGGATGGCCCTGACATTAATCTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGGCACTGCGGATGGCCCTGACATTAATCTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCCGACAAAGCGTTCCAGGACAACTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCCGACAAAGCGTTCCAGGACAACTGTAT 1800
QY 1801 CCGTTTACCTGGATGCGGTACGTTTAAACCGCAAGCTGATTGTTACCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGGTACGTTTAAACCGCAAGCTGATTGTTACCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCGAACCTGCTGCGACCGCCCAAAACCTGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCGAACCTGCTGCGACCGCCCAAAACCTGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTAAAGACGTTGGCGTGGATACGCTGGCGGAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAAGACGTTGGCGTGGATACGCTGGCGGAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAAACAAACACATGAATGAGACACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAAACAAACACATGAATGAGACACCGGATTAC 2160
QY 2161 TCCATCCGAGAGCTGCTTTTAAAGAGCGGAACACGCGATGACCATCAACGGCCGCTGG 2220
Db 2161 TCCATCCGAGAGCTGCTTTTAAAGAGCGGAACACGCGATGACCATCAACGGCCGCTGG 2220
QY 2221 GCATGGTCCACATCGACACAGCAAAAGTGAATTTATGTTGTAACGGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCACATCGACACAGCAAAAGTGAATTTATGTTGTAACGGTACTGCCGACCTTC 2280
QY 2281 AAGGTTCAACCATCCAAACCGTTGCTGGTGGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTTCAACCATCCAAACCGTTGCTGGTGGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400

Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCTGAAAACTATCTGCTGACTGATGAAGGCTCG 2400
QY 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCCGTAGCGTGAAGTCTTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCCGTAGCGTGAAGTCTTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCAAGTATTGGCGCCACCATGGAAAAACGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGCGAAAGATCCAAGTATTGGCGCCACCATGGAAAAACGCCAGAAAGGTGAAATCATG 2520
QY 2521 CCGAAACATCCGCGAGATGTCGGTCTTCTGCTGATGCGGTGCTACTCGGTGATCAACGCC 2580
Db 2521 CCGAAACATCCGCGAGATGTCGGTCTTCTGCTGATGCGGTGCTACTCGGTGATCAACGCC 2580
QY 2581 GCCAGCGGTCTGACAGCTGTCGATGAAGCCCTGAAAGACCGCAGACATAAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTCTGACAGCTGTCGATGAAGCCCTGAAAGACCGCAGACATAAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAAATTAACAATTAACAACCTCGGATCGAGGAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAAACAAATTAACAATTAACAACCTCGGATCGAGGAAGGATTCAGAAATTC 2700
QY 2701 GGATCTCTTCTCTGTCGGCCAGGGCGCTCGAGCCCGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTCGGCCAGGGCGCTCGAGCCCGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTGGTAAAGTCTTTCAGCGATCTCTGGCCACCTGCTTCGCCACACAGCGTACCCAC 2820
Db 2761 CCGGAATGTGGTAAAGTCTTTCAGCGAGCAGCTCTCCTGCTGGCCACACAGCGTACCCAC 2820
QY 2821 ACGGGTGAAAAACCGGTATTAATGCCAGATGCGGCAAAATCTTTTAGCACACAGCGGCTCC 2880
Db 2821 ACGGGTGAAAAACCGGTATTAATGCCAGATGCGGCAAAATCTTTTAGCACACAGCGGCTCC 2880
QY 2881 CTGGTGGCCCATCAACCGCATCTACTTGGCGAAGCCATACAAAATGTCAGAAATGTGGC 2940
Db 2881 CTGGTGGCCCATCAACCGCATCTACTTGGCGAAGCCATACAAAATGTCAGAAATGTGGC 2940
QY 2941 AAGTCTTTAGCCAGAGCTCCAGCTGGTGGCCACCAACGTACTCACACCGGGGAGAAG 3000
Db 2941 AAGTCTTTAGTGTGCTGATCTTGGAGGACCAACGTACTCACACCGGGGAGAAG 3000
QY 3001 CCTATGCTTTGTCGGAATGTGTAAAGTCTTCTCAGAGCTCTCACCTGGTGGCCAC 3060
Db 3001 CCTATGCTTTGTCGGAATGTGTAAAGTCTTCTCAGAGCTCTCACCTGGTGGCCAC 3060
QY 3061 CAGCGTACCCACACCGGTTGAAAAACCGTATAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACCGGTTGAAAAACCGTATAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
QY 3121 GACTGCGCGACCTTCTCGCCATCAACCGCATCTACTTGGCGAAGGACATCAAAATGT 3180
Db 3121 CGCAGGATAACCTTGTGCGCCATCAACCGCATCTACTTGGCGAAGGACATCAAAATGT 3180
QY 3181 CCAGAAATGTGGCAAGTCTTCTCCCAATCCAGCATCTCTGTCGGCACCACCAAGTACTCAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTCTCAACTTCAGGCCAATTTGGTCCGTCAACCAAGTACTCAC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCGCAGCGCCAGTACCCGTACGAGTTCGGGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCGCAGCGCCAGTACCCGTACGAGTTCGGGACTACGCT 3300

RESULT 6

AAH25681

ID AAH25681 standard; DNA; 8101 BP.

XX

AC AAH25681;

XX

DT 05-SEP-2001 (first entry)

XX

DE Nucleotide sequence of human Cdc25C phosphatase DNA.

XX Fusion protein; maltose binding protein; MBP; Cdc25C phosphatase;
KW Cdc25B1; Cdc25B2; Cdc25B3; cyclin-dependent kinase; CDK; cell division;
KW ss.
XX
XX Homo sapiens.
XX WO200144467-A2.
XX 21-JUN-2001.
XX
XX 13-DEC-2000; 2000WO-FR003496.
XX
XX 14-DEC-1999; 99FR-00015722.
XX 30-MAY-2000; 2000FR-00006883.
XX 21-SEP-2000; 2000FR-00012008.
XX
XX (SCRC) SCRAS SOC CONSEILS RECH & APPL SCI.
XX
XX Goubin-Gramatica F, Ducommun B, Prevost G;
XX WPI; 2001-398152/42.
XX
XX New fusion protein of maltose-binding protein and Cdc phosphatase, useful
XX for identifying phosphatase modulators for regulating the cell cycle.
XX
XX Claim 7; Page 9-13; 56pp; French.
XX
XX The present sequence encodes human Cdc25C phosphatase, and was expressed
XX in Escherichia coli. It is used to make fusion proteins with Escherichia
XX coli maltose binding protein (MBP). The specification describes fusion
XX proteins of MBP and one of Cdc 25B1, 25B2, 25B3 and 25C. Cdc phosphatases
XX are involved in activation of cyclin-dependent kinases (CDK) that control
XX cell division. The fusion proteins are used to identify modulators of the
XX specified human Cdc phosphatases, potentially useful for regulation of
XX cell division. They are also for studying physiological or
XX physiopathological activities of Cdc phosphatase
XX
XX Sequence 8101 BP; 2127 A; 1987 C; 2121 G; 1866 T; 0 U; 0 Other;
XX
XX
XX Query Match 82.0%; Score 2705.8; DB 4; Length 8101;
XX Best Local Similarity 99.9%; Pred. No. 0;
XX Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGGGTATGGCATGATAGCGCCGGAAGAGA 60
XX
XX 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGGGTATGGCATGATAGCGCCGGAAGAGA 60
XX
XX 61 GTCATTCAGGGTGGTGAATGTGAACCAAGTAACTTATACGATGTCGAGAGTATGCCG 120
XX
XX 61 GTCATTCAGGGTGGTGAATGTGAACCAAGTAACTTATACGATGTCGAGAGTATGCCG 120
XX
XX 121 GTGTCCTTTATCAGACCGTTTCCCGGTGGTGAACCAAGTAACTTATACGATGTCGAGAGT 180
XX
XX 121 GTGTCCTTTATCAGACCGTTTCCCGGTGGTGAACCAAGTAACTTATACGATGTCGAGAGT 180
XX
XX 181 CCGCGGAAAAAGTGAAGCGGCGATGGCGGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 240
XX
XX 181 CCGCGGAAAAAGTGAAGCGGCGATGGCGGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 240
XX
XX 241 AACAACTGGCGGCAAAAGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGCG 300
XX
XX 241 AACAACTGGCGGCAAAAGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGCG 300
XX
XX 301 ACGCGCGTGGCAAAATTTGGCGGCGATTAATCTCGCGCGCATCAATGGGTGGCAGCG 360
XX
XX 301 ACGCGCGTGGCAAAATTTGGCGGCGATTAATCTCGCGCGCATCAATGGGTGGCAGCG 360
XX
XX 361 TGGTGGTGTGATGTAGAAAGCGGCGTGAAGCCCTGTAAAGCGGCGGTGCAAAATC 420
XX
XX 361 TGGTGGTGTGATGTAGAAAGCGGCGTGAAGCCCTGTAAAGCGGCGGTGCAAAATC 420
XX
XX 421 TTCTCGCGCAACGGCTCAGTGGGTGATCATTAACCTATCCGCTGGATGACCAAGGATGCCA 480

Db
421 TTCTCGCGCAACGGCTCAGTGGGTGATCATTAACCTATCCCGCTGGATCACCAGGATGCCA 480
Qy
481 TTGCTGTGGAAGCTCCCTGCACCTAATGTTCCGGGGTTATTTCTTGATCTCTCTGACACAGA 540
Db
481 TTGCTGTGGAAGCTCCCTGCACCTAATGTTCCGGGGTTATTTCTTGATCTCTCTGACACAGA 540
Qy
541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC 600
Db
541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC 600
Qy
601 TGCTCGCATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCGTCTCGG 660
Db
601 TGCTCGCATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCGTCTCGG 660
Qy
661 CGCGTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGCGATAG 720
Db
661 CGCGTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGCGATAG 720
Qy
721 CGGAAACGGGAAAGCGCATGGAGTGCATGTCGGTTTTTCAACAAACCAATGCAAAATGCTGA 780
Db
721 CGGAAACGGGAAAGCGCATGGAGTGCATGTCGGTTTTTCAACAAACCAATGCAAAATGCTGA 780
Qy
781 ATGAGGCGATCGTTCCCACTCGGATGCTGGTGGCCAAACGATCAGATGGCGCTGGCGGCA 840
Db
781 ATGAGGCGATCGTTCCCACTCGGATGCTGGTGGCCAAACGATCAGATGGCGCTGGCGGCA 840
Qy
841 TGGCGCCATTTACCGAGTCCGGGTGCGGTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db
841 TGGCGCCATTTACCGAGTCCGGGTGCGGTGGTGGCGATATCTCGGTAGTGGGATACG 900
Qy
901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAAAACAGGATTTTC 960
Db
901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAAAACAGGATTTTC 960
Qy
961 GCCTGCTGGGCAAAACCGAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
Db
961 GCCTGCTGGGCAAAACCGAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
Qy
1021 AGGCAATCAGCTGTGTCCTCTCAGCTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
Db
1021 AGGCAATCAGCTGTGTCCTCTCAGCTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
Qy
1081 CGCAAAACCGCTCTCTCCCGCGTGGCGATTCATTAAATGCAAGTGGCAGACAGGTTT 1140
Db
1081 CGCAAAACCGCTCTCTCCCGCGTGGCGATTCATTAAATGCAAGTGGCAGACAGGTTT 1140
Qy
1141 CCGCACTGGAAGCGGCGAGTGAGGCGAACCGCAATTAATGTAGTTAGTCTCACTCATTTAG 1200
Db
1141 CCGCACTGGAAGCGGCGAGTGAGGCGAACCGCAATTAATGTAGTTAGTCTCACTCATTTAG 1200
Qy
1201 GCACAAATTCATGTTTGCACAGCTTATCATCGACTGCGGTCACCAATGCTTCTGGCG 1260
Db
1201 GCACAAATTCATGTTTGCACAGCTTATCATCGACTGCGGTCACCAATGCTTCTGGCG 1260
Qy
1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTGTAATCACTGCAATATTCG 1320
Db
1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTGTAATCACTGCAATATTCG 1320
Qy
1321 TGTGCTCAAGCGGCACTCCCGTCTCGATTAATGTTTTTGGCGCGACATCATACGGTT 1380
Db
1321 TGTGCTCAAGCGGCACTCCCGTCTCGATTAATGTTTTTGGCGCGACATCATACGGTT 1380
Qy
1381 CTGGCAAAATTTCTGAAATGAGCTGTTCGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
Db
1381 CTGGCAAAATTTCTGAAATGAGCTGTTCGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
Qy
1441 ATTGTGAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCAAGA 1500
Db
1441 ATTGTGAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCAAGA 1500
Qy
1501 GCATTTCACCAACAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAACTTCGG 1560

Db 1501 GCACCTTCACCAACAAGGACCATAGATTATGAAAAATCGAAGAGGTAAATCTGGTAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCGAAGAGAAATTCAGAGAAATTC 1680
Db 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCGAAGAGAAATTCAGAGAAATTC 1680
QY 1681 GGGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACAGCCGCTTTGGTGGCTAC 1740
Db 1681 GGGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACAGCCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTGGCTGAAATCAACCCGGAGAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTGGCTGAAATCAACCCGGAGAAAGCGTTCCAGGACAAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTACCAACGCAAGCTGATTGCTTACCAGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACCAACGCAAGCTGATTGCTTACCAGATCGCTGTT 1860
QY 1861 GAAGGTTTATCGCTGATTATTAACAAGATCTGCTGCGAACCCGCCAAACCTGGAA 1920
Db 1861 GAAGGTTTATCGCTGATTATTAACAAGATCTGCTGCGAACCCGCCAAACCTGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATTAAGAAGTCAAGCGAAGAGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATTAAGAAGTCAAGCGAAGAGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATTGCTGCTGACGGGGGTTATGCTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATTGCTGCTGACGGGGGTTATGCTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATACGCTGGCGGCAAGAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATACGCTGGCGGCAAGAGCG 2100
QY 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAACCAACACACATGAATGCAGACACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAACCAACACACATGAATGCAGACACCGGATTAC 2160
QY 2161 TCCATCCGAGAGCTGCTTTAATAAGGCGAACAACAGGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCCGAGAGCTGCTTTAATAAGGCGAACAACAGGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTAATGCTGTAACGGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTAATGCTGTAACGGTACTGCCGACCTTC 2280
QY 2281 AAGGGTCAACATCCAAACCGTTCGTTGGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGGTCAACATCCAAACCGTTCGTTGGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTCTCTCGAAACTATCTGCTGACTGATGAGGCTCG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTCTCTCGAAACTATCTGCTGACTGATGAGGCTCG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCTGTAAGTCTTACGAGGAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCTGTAAGTCTTACGAGGAGAG 2460
QY 2461 TTGGCGAAGATCCACGTTATGCGGCACCATGGAACACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGTTATGCGGCACCATGGAACACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGAGATGTCGCTTTCTGTTATGCGGTGCTGCTGCGGTGATCAAGCC 2580
Db 2521 CCGAACATCCCGAGATGTCGCTTTCTGTTATGCGGTGCTGCTGCGGTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGACACTGCTGATGAAGCCCTGAAAGACCGCGAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACACTGCTGATGAAGCCCTGAAAGACCGCGAGACTAATTCGAGCTCG 2640

QY 2641 AACCAACACACATTAACAATAACAACACCTCGGATCGAGGAAGATTTCAGAAATTC 2700
Db 2641 AACCAACACACATTAACAATAACAACACCTCGGATCGAGGAAGATTTCAGAAATTC 2700
QY 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709

RESULT 7
ABL53238
ID ABL53238 standard; DNA; 8101 BP.
XX ABL53238;
AC ABL53238;
XX
DT 17-JUN-2002 (first entry)
XX
DE Nucleotide sequence of pMAL-Hs Cdc25C.
XX
KW Cdc25C phosphatase; amine; Cdc25 phosphatase; spontaneous alopecia;
KW alopecia; proliferative disease; parasitic disease; viral infection;
KW neurodegenerative disease; myopathy; angiogenesis; psoriasis; restenosis;
KW vitamin K; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1528..2677
FT /*tag= b
FT /note= "encodes maltose binding protein (MBP)"
FT CDS 2713..4134
FT /*tag= a
FT /note= "ORF of human Cdc25C"
XX
PN FR2812198-A1.
XX
PD 01-FEB-2002.
XX
PF 28-JUL-2000; 2000FR-00009900.
XX
PR 28-JUL-2000; 2000FR-00009900.
XX
PA (SCRC) SCRAS SOC CONSEILS RECH & APPL SCI.
XX
PI Prevost G, Brezak PMC, Galcera CMO, Thureau C, Goubin GF;
PI Ducommun B;
XX
DR WPI; 2002-166066/22.
XX
XX
PT cdc25 Phosphatase inhibiting medicaments, useful e.g. for treating tumor
PT diseases, viral infections, neurodegenerative disease or alopecia,
PT containing new or known aryl-substituted amines.
XX
PS Disclosure; Page 18-22; 39pp; French.
XX
CC The present sequence represents pMAL-HsCdc25C, a plasmid encoding human
CC Cdc25C phosphatase. The specification describes the use of aryl-
CC substituted secondary or tertiary amines for the production of
CC medicaments for inhibiting Cdc25 phosphatases, especially Cdc25C
CC phosphatases. The amines of the invention are used for treating
CC spontaneous alopecia or alopecia induced by exogenous products or
CC radiation. They are also used for treating tumoral or non-tumoral
CC proliferative diseases, parasitic diseases, viral infections,
CC neurodegenerative diseases or myopathy. Examples of non-tumoral
CC proliferative diseases are e.g. angiogenesis, psoriasis or restenosis.
CC The amines are also useful for inhibiting the proliferation of
CC microorganisms (especially yeasts) and in the treatment of all disorders
CC conventionally treated with vitamin K or its derivatives
XX
SQ Sequence 8101 BP; 2127 A; 1987 C; 2121 G; 1866 T; 0 U; 0 Other;

Query Match 82.0%; Score 2705.8; DB 6; Length 8101; Best Local Similarity 99.9%; Pred. No. 0; Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;									
QY	1	CGGACACCATCAATGGTGCAGAACTCTTCGGCGGTATGGCATGATAGCCCGGAAGAGA	60						
DB	1	CGGACACCATCAATGGTGCAGAACTCTTCGGCGGTATGGCATGATAGCCCGGAAGAGA	60						
QY	61	GTCAATTCAGGTGTGAATGTGAACCAAGTAACTGTTATACGATGTGCGAGATATGCGG	120						
DB	61	GTCAATTCAGGTGTGAATGTGAACCAAGTAACTGTTATACGATGTGCGAGATATGCGG	120						
QY	121	GTGTCTCTTATCAGACCGTTTCCCGGTGTGAACCAAGTAACTGTTATACGATGTGCGAG	180						
DB	121	GTGTCTCTTATCAGACCGTTTCCCGGTGTGAACCAAGTAACTGTTATACGATGTGCGAG	180						
QY	181	CGCGGGAAGGTGAAGCGCGATGGCGGATGAATTAATTAATTAATTAATTAATTAATTA	240						
DB	181	CGCGGGAAGGTGAAGCGCGATGGCGGATGAATTAATTAATTAATTAATTAATTAATTA	240						
QY	241	AAACAATGCGGGCAACAGTCTGTGCTGATTTGGCGTTGCCACCTCTGAGTCTGGCCCTGC	300						
DB	241	AAACAATGCGGGCAACAGTCTGTGCTGATTTGGCGTTGCCACCTCTGAGTCTGGCCCTGC	300						
QY	301	ACGCGCGTGCAGAAATTTGTCGGCGGATTAATTAATTAATTAATTAATTAATTAATTA	360						
DB	301	ACGCGCGTGCAGAAATTTGTCGGCGGATTAATTAATTAATTAATTAATTAATTAATTA	360						
QY	361	TGTTGTTGTCGATGTTAGAACGAAGCGCGTGAAGCCTGTAAGCGGCGTGCAATC	420						
DB	361	TGTTGTTGTCGATGTTAGAACGAAGCGCGTGAAGCCTGTAAGCGGCGTGCAATC	420						
QY	421	TTTCTCGGCAACCGCTCAGTGGGCTGATCAATTAATTAATTAATTAATTAATTAATTA	480						
DB	421	TTTCTCGGCAACCGCTCAGTGGGCTGATCAATTAATTAATTAATTAATTAATTAATTA	480						
QY	481	TTGCTGTGGAAGCTGCTGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTA	540						
DB	481	TTGCTGTGGAAGCTGCTGCTCAATTAATTAATTAATTAATTAATTAATTAATTAATTA	540						
QY	541	CACCCATCAACAGTATATTTTCTCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC	600						
DB	541	CACCCATCAACAGTATATTTTCTCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC	600						
QY	601	TGTTGCGCATTTGGTCAACAGCAAACTGCGCTGTTAGCGGCGCCATTAAGTCTCTCGG	660						
DB	601	TGTTGCGCATTTGGTCAACAGCAAACTGCGCTGTTAGCGGCGCCATTAAGTCTCTCGG	660						
QY	661	CGGCTGCGGCTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
DB	661	CGGCTGCGGCTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
QY	721	CGGAACGGGAAGCGGACTGAGTGCCATCTCGGTTTTCAACAAACCATGCAAAATGCTGA	780						
DB	721	CGGAACGGGAAGCGGACTGAGTGCCATCTCGGTTTTCAACAAACCATGCAAAATGCTGA	780						
QY	781	ATGAGGGCATCTGTTCCCACTGCGATGCTGTTGCGCAACGATCAGATGGCGTGGCGCAA	840						
DB	781	ATGAGGGCATCTGTTCCCACTGCGATGCTGTTGCGCAACGATCAGATGGCGTGGCGCAA	840						
QY	841	TGCGGCGCATACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG	900						
DB	841	TGCGGCGCATACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG	900						
QY	901	ACGATACCAAGACAGCTCATGTTATATCCCGCGTTTAAACCAATCAACAGGATTTTC	960						
DB	901	ACGATACCAAGACAGCTCATGTTATATCCCGCGTTTAAACCAATCAACAGGATTTTC	960						
QY	961	GCTGTGCGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGGCGTGA	1020						
DB	961	GCTGTGCGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGGCGTGA	1020						
QY	1021	AGGGCAATCAGTGTGTTGCCCGTCTCACTGGTGAAGAAACCAACCTGGCGGCCCAATA	1080						

DB	1021	AGGGCAATCAGTGTGTTGCCCGTCTCACTGGTGAAGAAACCAACCTGGCGGCCAATA	1080						
QY	1081	CGCAAAACCGCCTCTCCCGCGCGTTGGCGGATTCATTAAATGAGCTGGCACGACAGGTTT	1140						
DB	1081	CGCAAAACCGCCTCTCCCGCGCGTTGGCGGATTCATTAAATGAGCTGGCACGACAGGTTT	1140						
QY	1141	CCGACTGGAAGCGGCGAGTGAGGCAACGCAATTAATGTGAGTTAGTCTACTCATTAG	1200						
DB	1141	CCGACTGGAAGCGGCGAGTGAGGCAACGCAATTAATGTGAGTTAGTCTACTCATTAG	1200						
QY	1201	GCACAAATTCATGTTTCAGACGCTTATCATCGACTGACGCTGCACCAATGCTTCTGGCG	1260						
DB	1201	GCACAAATTCATGTTTCAGACGCTTATCATCGACTGACGCTGCACCAATGCTTCTGGCG	1260						
QY	1261	TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTTCGTAATCACTGCATTAATTCG	1320						
DB	1261	TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTTCGTAATCACTGCATTAATTCG	1320						
QY	1321	TGTCGCTCAAGCGGCACTCCCGTTCTGGATTAATGTTTTTGGCGCGGACATCAATAACGGTT	1380						
DB	1321	TGTCGCTCAAGCGGCACTCCCGTTCTGGATTAATGTTTTTGGCGCGGACATCAATAACGGTT	1380						
QY	1381	CTGGCAAAATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATCACTGATGTGA	1440						
DB	1381	CTGGCAAAATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATCACTGATGTGA	1440						
QY	1441	ATTGTGAGCGGATTAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAGTGTTTTTTCACGA	1500						
DB	1441	ATTGTGAGCGGATTAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAGTGTTTTTTCACGA	1500						
QY	1501	GCACCTTCCAAACGAAGCCATAGATTATGAAAACTGAAGAAAGTAAACTGGTAAATCTGG	1560						
DB	1501	GCACCTTCCAAACGAAGCCATAGATTATGAAAACTGAAGAAAGTAAACTGGTAAATCTGG	1560						
QY	1561	ATTAAACGGGATTAAGGCTTAAACGGCTCTCGCTGAAGTTCGTAAGTTCGTAAGTTCG	1620						
DB	1561	ATTAAACGGGATTAAGGCTTAAACGGCTCTCGCTGAAGTTCGTAAGTTCGTAAGTTCG	1620						
QY	1621	ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACACAGGTT	1680						
DB	1621	ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACACAGGTT	1680						
QY	1681	GGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACGACGCTTGGTGGCTAC	1740						
DB	1681	GGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACGACGCTTGGTGGCTAC	1740						
QY	1741	GCTCAATCTGGCCTGTTGGCTGAAATCACCCGGACAAAGCGTTCCAGGACAAAGCTGTAT	1800						
DB	1741	GCTCAATCTGGCCTGTTGGCTGAAATCACCCGGACAAAGCGTTCCAGGACAAAGCTGTAT	1800						
QY	1801	CGGTTTACCTGGGATGCGGTACGTTTCAACCGCAAGCTGATTGCTTACCCGATCGCTGTT	1860						
DB	1801	CGGTTTACCTGGGATGCGGTACGTTTCAACCGCAAGCTGATTGCTTACCCGATCGCTGTT	1860						
QY	1861	GAAAGCGTTATCGCTGATTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA	1920						
DB	1861	GAAAGCGTTATCGCTGATTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA	1920						
QY	1921	GAGATCCCGCGCTGGATAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTCAAC	1980						
DB	1921	GAGATCCCGCGCTGGATAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTCAAC	1980						
QY	1981	CTGCAAGAAACCGTACTTCACTCGGCGCTGATTGCTGCTGACGGGGTTATCGCTTCAAG	2040						
DB	1981	CTGCAAGAAACCGTACTTCACTCGGCGCTGATTGCTGCTGACGGGGTTATCGCTTCAAG	2040						
QY	2041	TATGAAAAACCGCAAGTACGACATTAAGACGTTGGCGCTGGATTAACGCTGGCGGAAAGCG	2100						
DB	2041	TATGAAAAACCGCAAGTACGACATTAAGACGTTGGCGCTGGATTAACGCTGGCGGAAAGCG	2100						
QY	2101	GCTCTGACCTTCTGTTGACCTGATTAATAACCAACATGAATGCAGACACCGATAC	2160						

```
Db 2101 GGTCTGACCTTCTGTGTTGACCTGATTAAAAACAACACATGAATGACGACACCGATTAC 2160
QY 2161 TCCATCGCAGAAGCTGCTTTTAAATAAGCGGAAACAGCGATGACCATCAACGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAAGCTGCTTTTAAATAAGCGGAAACAGCGATGACCATCAACGCGCCGCTGG 2220
QY 2221 GCATGGTCAACATCGACACGACGAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACGACGAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTTGTGTCGTCGTGAGCGCAGGTATTAAAGCGCCCGAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTTGTGTCGTCGTGAGCGCAGGTATTAAAGCGCCCGAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGATTGCGCCACCATGGAACGCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTGCGCCACCATGGAACGCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAATCATCCGCGAGATGTCGCTTTTCTGGTATGCGGTGCTGCTGATCAACGCC 2580
Db 2521 CCGAATCATCCGCGAGATGTCGCTTTTCTGGTATGCGGTGCTGCTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGCGCGCAGACTAATTCAGCTCG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGCGCGCAGACTAATTCAGCTCG 2640
QY 2641 AACAAACAACAATAACAATAACAACCAACCTCGGGATCGAGGGAAGATTTCAGAAATC 2700
Db 2641 AACAAACAACAATAACAATAACAACCAACCTCGGGATCGAGGGAAGATTTCAGAAATC 2700
QY 2701 GGATCCCTCT 2709
Db 2701 GGATCCCTCT 2709

RESULT 8
AD023608
ID AD023608 standard; DNA; 7259 BP.
AC
XX
AC AD023608;
XX
DT 01-JUL-2004 (first entry)
XX
DE DNA encoding MBP-Toxop30del18 (82-294aa) fusion protein.
XX
KW P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; de; gene.
XX
OS Toxoplasma gondii.
OS Synthetic.
XX
FN US2004067239-A1.
XX
PD 08-APR-2004.
XX
PF 02-OCT-2002; 2002US-00263153.
XX
PR 02-OCT-2002; 2002US-00263153.
XX
PA (MAIN/) MAINE G T.
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX
```

PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;

XX WPI; 2004-304563/28.

DR P-PSDB; ADO23609.

XX Novel purified polypeptide having sequence identity to amino acid

PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,

PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of Igm or IgG

PT antibodies to Toxoplasma gondii.

XX Example 2; Fig 22; 114pp; English.

PS The invention relates to a purified P30 antigen (I) chosen from 3 fully

XX defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6

CC amino acids added to the C-terminus of the amino acid sequence of

CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in

CC which at least one of the five C-terminal cysteine amino acids of the

CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted

CC with alanine, or comprising the amino acid sequence chosen from MBP-

CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1

CC fusion proteins. (I) is useful for detecting the presence of Igm

CC antibodies to Toxoplasma gondii in a test sample, which involves

CC contacting the test sample suspected of containing the Igm antibodies

CC with a composition comprising (I) and detecting the presence of (I)/Igm

CC antibody complexes. The present sequence represents DNA encoding a MBP-

CC ToxoP30 fusion protein of the invention.

XX Sequence 7259 BP; 1870 A; 1822 C; 1926 G; 1641 T; 0 U; 0 Other;

QY Query Match 81.5%; Score 2690.4; DB 12; Length 7259;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

1 CCGACACCATCGATGGTGC AAAACCTTTCCGGGTATGGCATGATAGCGCCCGAAGAGA 60

1 CCGACACCATCGATGGTGC AAAACCTTTCCGGGTATGGCATGATAGCGCCCGAAGAGA 60

61 GTCAATTCAAGGTTGGTGAATGTGAACACAGTACGTTATACGATGTCGAGATATGCCG 120

61 GTCAATTCAAGGTTGGTGAATGTGAACACAGTACGTTATACGATGTCGAGATATGCCG 120

121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACCCAGCCAGCCACGTTTCTGCGAAAA 180

121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACCCAGCCAGCCACGTTTCTGCGAAAA 180

181 CGCGGAAAAAAGTGAAGCGCGGATGGCGAGCTGAATTACATTCACACCCGCTGGCAC 240

181 CGCGGAAAAAAGTGAAGCGCGGATGGCGAGCTGAATTACATTCACACCCGCTGGCAC 240

241 AACAACTGGCGGGCAACAGTCTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300

241 AACAACTGGCGGGCAACAGTCTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300

301 AGCGCGCGTGGCAAAATGTCGGCGGATTAATCTCGGCCGATCACTGGGTGCGCAGCG 360

301 AGCGCGCGTGGCAAAATGTCGGCGGATTAATCTCGGCCGATCACTGGGTGCGCAGCG 360

361 TGGTGTGTGATGGTAGAACGAGCGGCTCGAAGCCCTGTAAAGCGGCTGCAATC 420

361 TGGTGTGTGATGGTAGAACGAGCGGCTCGAAGCCCTGTAAAGCGGCTGCAATC 420

421 TTCTCGCGCAACCGCTCAGTGGGTGATCACTAATTAATCTCCGCTGGATGACACAGGATGCCA 480

421 TTCTCGCGCAACCGCTCAGTGGGTGATCACTAATTAATCTCCGCTGGATGACACAGGATGCCA 480

481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCTTATTTCTGATGTCTTGACACAGA 540

481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCTTATTTCTGATGTCTTGACACAGA 540

541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600

541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600

QY 601 TGGTCGCAATGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
DB 601 TGGTCGCAATGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGCTCGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTCGCTCGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGCGCATGGAGTGCATGTCGGTTCATCAAAACCATGCAAAATGCTGA 780
DB 721 CGGAACGGGAAGCGCATGGAGTGCATGTCGGTTCATCAAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGCGATCGTCCCACTCGCATGCTGGTTGCCAACGATCGAGTGGCGTGGCGGCA 840
DB 781 ATGAGGCGATCGTCCCACTCGCATGCTGGTTGCCAACGATCGAGTGGCGTGGCGGCA 840
QY 841 TCGCGGCCATTAACCGAGTCCGGCTCGCGTGGTGGGATATCTCGGTAGTGGGATAG 900
DB 841 TCGCGGCCATTAACCGAGTCCGGCTCGCGTGGTGGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACCATCAAAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACCATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
DB 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACACCTGGCGGCCAATA 1080
DB 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACACCTGGCGGCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCACTGGCGACACAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCACTGGCGACACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGGCGCAACGCAATTAATGTCAGTTAGTCACTCATTAG 1200
DB 1141 CCCGACTGGAAGCGGCGAGTGGCGCAACGCAATTAATGTCAGTTAGTCACTCATTAG 1200
QY 1201 GCACAATCTCATGTTTGACAGCTTATCATCGACTGACGGTGCACCAATGCTCTGGCG 1260
DB 1201 GCACAATCTCATGTTTGACAGCTTATCATCGACTGACGGTGCACCAATGCTCTGGCG 1260
QY 1261 TCAGGCGAGCATCGGAAGCTGTGGTATGCTGTGCAGGTGTAATCACTGCATTAATTCG 1320
DB 1261 TCAGGCGAGCATCGGAAGCTGTGGTATGCTGTGCAGGTGTAATCACTGCATTAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAATAACGGTT 1380
DB 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAATAACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAATTAATCACTCGCTGTAATGTTGGA 1440
DB 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAATTAATCACTCGCTGTAATGTTGGA 1440
QY 1441 ATTGTGAGCGGATAACAAATTCACACAGGAACAGCCAGTCCGTTTAGTGTTTCACGA 1500
DB 1441 ATTGTGAGCGGATAACAAATTCACACAGGAACAGCCAGTCCGTTTAGTGTTTCACGA 1500
QY 1501 GCATTTACCAACAGGACATAGATATGAAAACTGAAAGAGGTAAACTGTAATCTCG 1560
DB 1501 GCATTTACCAACAGGACATAGATATGAAAACTGAAAGAGGTAAACTGTAATCTCG 1560
QY 1561 ATTAACGGCGATAAAGGCTATACCGTCTCGCTGAGTGGTAAGAAATTCAGAAAGAT 1620
DB 1561 ATTAACGGCGATAAAGGCTATACCGTCTCGCTGAGTGGTAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAATTTAAAGTTCACCGCTGAGCATCCGGATTAACCTGGAAGAGAAATTCACAGGTT 1680
DB 1621 ACCGGAATTTAAAGTTCACCGCTGAGCATCCGGATTAACCTGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGGCCCTGACATTAATCTCTGGGCAACAGACCGCTTTGGTGGCTAC 1740

RESULT 9
AD023613
ID AD023613 standard; DNA; 7322 BP.
XX AC AD023613;

DB 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCAACAGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCCGAACAAGCGTTCCAGGACAAGCTGTAT 1800
DB 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCCGAACAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTACAAACGGCAAGCTGATTGCTTACCCGATCGCTGT 1860
DB 1801 CCGTTTACCTGGGATGCGGTACGTTACAAACGGCAAGCTGATTGCTTACCCGATCGCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
DB 1861 GAAGCGTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAACGAAAGTAAAGAGCGCTGATGTTCAAC 1980
DB 1921 GAGATCCCGCGCTGGATTAAGAACTGAAACGAAAGTAAAGAGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
DB 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACGCGCAAGTACGACATTAAGACCGTGGCGTGGATAACGCTGGCGGAAAGCG 2100
DB 2041 TATGAAACGCGCAAGTACGACATTAAGACCGTGGCGTGGATAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACACATGAATGCAACACCGATTAC 2160
DB 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACACATGAATGCAACACCGATTAC 2160
QY 2161 TCCATCGAGAAGCTGCGCTTTAATAAAGCGAACAAGCATGACCATCAACCGCCGCTGG 2220
DB 2161 TCCATCGAGAAGCTGCGCTTTAATAAAGCGAACAAGCATGACCATCAACCGCCGCTGG 2220
QY 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAATGTTGTAACCGTACTGCGGACCTTC 2280
DB 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAATGTTGTAACCGTACTGCGGACCTTC 2280
QY 2281 AAGGTCACCAATCAAAACCGTTCTGGCGTGTGAGCGCAGGTATTAAACCGCCCACT 2340
DB 2281 AAGGTCACCAATCAAAACCGTTCTGGCGTGTGAGCGCAGGTATTAAACCGCCCACT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAATACTATCTGCTGACTGATGAAGGTCTG 2400
DB 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAATACTATCTGCTGACTGATGAAGGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTCAAGTCTTACGAGGAAG 2460
DB 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTCAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCAAAAGATCCACGTTATTCGCCCACTTGGAAAAACGCCAGAAAGGTGAAATCATG 2520
DB 2461 TTGGCAAAAGATCCACGTTATTCGCCCACTTGGAAAAACGCCAGAAAGGTGAAATCATG 2520
QY 2521 CCGAACAATCCCGAGATGTCGCTTTTGGTATGCGGTGCTGCTGCTGCTGCTGCTGCTG 2580
DB 2521 CCGAACAATCCCGAGATGTCGCTTTTCTGGTATGCGGTGCTGCTGCTGCTGCTGCTG 2580
QY 2581 GCCAGCGCTGCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTG 2640
DB 2581 GCCAGCGCTGCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTG 2640
QY 2641 AACCAACAACAATAAACAACAACCAACCTCGGGATCGAGGGAGGATTTTCAGAAATTC 2700
DB 2641 AACCAACAACAATAAACAACAACCAACCTCGGGATCGAGGGAGGATTTTCAGAAATTC 2700

XX
DT 01-JUL-2004 (first entry)
XX DNA encoding MBP-Toxop30del110 (52-284aa) fusion protein.
XX DE
XX P30 antigen; Toxop30del13C; Toxop30del12C; Toxop30 MIX1;
KW MBP-Toxop30del13C(52-300aa); MBP-Toxop30del14C(52-294aa); MBP-Toxop30MIX1;
KW Toxoplasma gondii; de; gene.
XX
OS Toxoplasma gondii.
OS Synthetic.
XX
XX US2004067239-A1.
XX
XX 08-APR-2004.
XX
XX 02-OCT-2002; 2002US-00263153.
XX
XX 02-OCT-2002; 2002US-00263153.
XX
XX (MAIN/) MAINE G T.
XX (PATE/) PATEL C B.
XX (GINS/) GINSBURG S R.
XX (BLIE/) BLIESE T R.
XX
XX Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX WPI; 2004-304563/28.
XX P-PSDB; ADO23614.
XX
XX Novel purified polypeptide having sequence identity to amino acid
XX sequence of p30 antigen of Toxoplasma gondii e.g. Toxop30del13C,
XX Toxop30del12C, Toxop30 MIX1, useful for detecting presence of IgM or IgG
XX antibodies to Toxoplasma gondii.
XX
XX Example 2; Fig 25; 114pp; English.
XX
XX The invention relates to a purified p30 antigen (I) chosen from 3 fully
XX defined Toxop30del13C, Toxop30del12C and Toxop30 MIX1 sequences, having 1-6
XX amino acids added to the C-terminus of the amino acid sequence of
XX Toxop30del12C p30 antigen sequence, or comprising an amino acid sequence in
XX which at least one of the five C-terminal cysteine amino acids of the
XX amino acid sequence of Toxop30del13C p30 antigen sequence is substituted
XX with alanine, or comprising the amino acid sequence chosen from MBP-
XX Toxop30del13C(52-300aa), MBP-Toxop30del14C(52-294aa) and MBP-Toxop30MIX1
XX fusion proteins. (I) is useful for detecting the presence of IgM
XX antibodies to Toxoplasma gondii in a test sample, which involves
XX contacting the test sample suspected of containing the IgM antibodies
XX with a composition comprising (I) and detecting the presence of (I)/IgM
XX antibody complexes. The present sequence represents DNA encoding a MBP-
XX Toxop30 fusion protein of the invention.
XX
XX Sequence 7322 BP; 1887 A; 1844 C; 1934 G; 1657 T; 0 U; 0 Other;
XX
XX Query Match 81.5%; Score 2690.4; DB 12; Length 7322;
XX Best Local Similarity 99.8%; Pred. No. 0;
XX Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
XX
XX 1 CCGACACATCGAATGTGCAAAACCTTTCGCGTATGGCATGATAGCGCCCGAAGAGA 60
XX
XX 1 CCGACACATCGAATGTGCAAAACCTTTCGCGTATGGCATGATAGCGCCCGAAGAGA 60
XX
XX 61 GTCAATTCAGGGTGGTGAATGTGAACACAGTAAAGTTATACGATGTCGACAGATATGCCG 120
XX
XX 61 GTCAATTCAGGGTGGTGAATGTGAACACAGTAAAGTTATACGATGTCGACAGATATGCCG 120
XX
XX 121 GTGTCTCTTATCAGACCGTTTCCGCGTGGTGAACACGAGCCAGCCAGTTTCTGCGAAAA 180
XX
XX 121 GTGTCTCTTATCAGACCGTTTCCGCGTGGTGAACACGAGCCAGCCAGTTTCTGCGAAAA 180
XX
XX 181 CGCGGGAAAGTGGAGCGCGGATGCGGAGCTGAATTCATATCCCAACCGCGTGGCAC 240
XX
XX 181 CGCGGGAAAGTGGAGCGCGGATGCGGAGCTGAATTCATATCCCAACCGCGTGGCAC 240

QY 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATGCGGTGCGACCTCCAGTCTGCCCCTGC 300
DB 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGCGTGTGCACCTCCAGTCTGCCCCTGC 300
QY 301 ACGCGCGTCCCAAAATTTGTCGCGCGATTAATCTCGCGCGGATCAACTGGGTGCGACGC 360
DB 301 ACGCGCGTCCCAAAATTTGTCGCGCGATTAATCTCGCGCGGATCAACTGGGTGCGACGC 360
QY 361 TGGTGGTGTGATGGTAGAAGCGGCGTGAAGCGCTGTAAAGCGGCGGTGCACAATC 420
DB 361 TGGTGGTGTGATGGTAGAAGCGGCGTGAAGCGCTGTAAAGCGGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCACTCCGCTGGATGACACAGATGCCA 480
DB 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCACTCCGCTGGATGACACAGATGCCA 480
QY 481 TTGCTGTGAAGCTGCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
DB 481 TTGCTGTGAAGCTGCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
QY 541 CACCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 600
DB 541 CACCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 600
QY 601 TGGTCGCATTTGGGTACCCAGCAATCGCGCTGTAGCGGGCCCAATTAAGTCTCTCGG 660
DB 601 TGGTCGCATTTGGGTACCCAGCAATCGCGCTGTAGCGGGCCCAATTAAGTCTCTCGG 660
QY 661 CGGCTGTGCTGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGGCTGTGCTGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGCGGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
DB 721 CGGAACGGGAAGCGGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGCGCATCGTTCCCACTCGCATGCTGTTCCCAACCATGATGCGCTGGCGGCA 840
DB 781 ATGAGGCGCATCGTTCCCACTCGCATGCTGTTCCCAACCATGATGCGCTGGCGGCA 840
QY 841 TGGCGCGCATTAACGAGTCCGGCTCGCGGTGGTGGCGGATATCTCGTAGTGGGATAG 900
DB 841 TGGCGCGCATTAACGAGTCCGGCTCGCGGTGGTGGCGGATATCTCGTAGTGGGATAG 900
QY 901 ACATACCGAAGACGCTCATGTTATATCCCGGTAAACCAACCATCAACAGGATTTTC 960
DB 901 ACATACCGAAGACGCTCATGTTATATCCCGGTAAACCAACCATCAACAGGATTTTC 960
QY 961 GCCTGTGGGGCAAAACCGGCTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGGCGTGA 1020
DB 961 GCCTGTGGGGCAAAACCGGCTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGGCGTGA 1020
QY 1021 AGGGCAATCAGCTGTGCGCGTCTCACTGGTGAAGAAAAACCAACCTTGGCGGCCCAATA 1080
DB 1021 AGGGCAATCAGCTGTGCGCGTCTCACTGGTGAAGAAAAACCAACCTTGGCGGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGCGTTGGCGCATTCATTATGAGCTGGCAGCTGGCACAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGCGTTGGCGCATTCATTATGAGCTGGCAGCTGGCACAGGTTT 1140
QY 1141 CCGGACTGAAACCGGCGCATGAGCGCAACGCAATTAATGTAGTTAGTCTCACTCATTAG 1200
DB 1141 CCGGACTGAAACCGGCGCATGAGCGCAACGCAATTAATGTAGTTAGTCTCACTCATTAG 1200
QY 1201 GCACAATTCATGTTTGACAGCTTATCATCGACTGCGCGGTGACCAATGCTTCTGGCG 1260
DB 1201 GCACAATTCATGTTTGACAGCTTATCATCGACTGCGCGGTGACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACGACCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTAAATCACTGCAATATTCG 1320
DB 1261 TCAGGACGACCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTAAATCACTGCAATATTCG 1320

QY 1321 TGTCTCAAGCGGCACATCCCGCTTCGGATATATGTTTTTGGCCGACATCATACGGTT 1380
Db 1321 TGTCTCAAGCGGCACATCCCGCTTCGGATATATGTTTTTGGCCGACATCATACGGTT 1380
QY 1381 CTGGCAATATCTGAATAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAATATCTGAATAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
QY 1441 ATTGTGAGCGGATACAAATTTTACACAGGAAACAGCCAGTCCGTTTGTAGTGTTCACGA 1500
Db 1441 ATTGTGAGCGGATACAAATTTTACACAGGAAACAGCCAGTCCGTTTGTAGTGTTCACGA 1500
QY 1501 GCATTTCAACCAAGGACCATAGATATGAAAATCGAAGAGGTAAATCTGGTAATCTGG 1560
Db 1501 GCATTTCAACCAAGGACCATAGATATGAAAATCGAAGAGGTAAATCTGGTAATCTGG 1560
QY 1561 ATTTAAGCGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTTAAGCGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGAAGAGAAATTTCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGAAGAGAAATTTCCACAGGTT 1680
QY 1681 GCGGCAATCGGCGATGCGCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAATCGGCGATGCGCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGCTGGCTGAATATCAACCGCAAGCTGATTCCTTACCCGATCCCTGTT 1800
Db 1741 GCTCAATCTGCGCTGCTGGCTGAATATCAACCGCAAGCTGATTCCTTACCCGATCCCTGTT 1800
QY 1801 CCGTTTACCTGGATGCGCTGATTAACACGCAAGCTGATTCCTTACCCGATCCCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGCTGATTAACACGCAAGCTGATTCCTTACCCGATCCCTGTT 1860
QY 1861 GAAGGTTATCGCTGATTTAACAAGATCTGCTGCCGAACCCCGCAAAAACCTGGGAA 1920
Db 1861 GAAGGTTATCGCTGATTTAACAAGATCTGCTGCCGAACCCCGCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGATAAAGAACTGAAGCGAAGGTAAAGCGCGTGAATTCACAC 1980
Db 1921 GAGATCCCGGCGCTGATAAAGAACTGAAGCGAAGGTAAAGCGCGTGAATTCACAC 1980
QY 1981 CTGCAAGAACCGTACTTACCTGCGCGCTGATTCCTGCTGACGGGGTTATCGCTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTGCGCGCTGATTCCTGCTGACGGGGTTATCGCTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGTTGGCGGTGATTAACGTCGCGCAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGTTGGCGGTGATTAACGTCGCGCAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAACCAACACATCAATGCACACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAACCAACACATCAATGCACACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAAGCGCAACAGCGATCACCATCAACGGCCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAAGCGCAACAGCGATCACCATCAACGGCCCGCTGG 2220
QY 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTTATGGTGAACGTTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTTATGGTGAACGTTACTGCGGACCTTC 2280
QY 2281 AAGGTTCAACCATCAACCGTTGTTGGCGTGTGAGCGGAGGTATTAACCGCGCAGT 2340
Db 2281 AAGGTTCAACCATCAACCGTTGTTGGCGTGTGAGCGGAGGTATTAACCGCGCAGT 2340
QY 2341 CCGAACAAGAGCTGCGCAAGAGTTCCTCGAAAACTATCTGCTGACTGATCAAGTCTG 2400
Db 2341 CCGAACAAGAGCTGCGCAAGAGTTCCTCGAAAACTATCTGCTGACTGATCAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGCTAGCGCTGAAGTCTTACGAGGAAG 2460

Db 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGCTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGCGAAAGATCCACGATTTTGGCGCCACCATGGAACACGCCAGAAAAGGTGAATCATG 2520
Db 2461 TTGCGAAAGATCCACGATTTTGGCGCCACCATGGAACACGCCAGAAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGGCTTTCTGGTATGTCGGTGCCTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGGCTTTCTGGTATGTCGGTGCCTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGGTGCTGACACTGTCGATGAAGCCCTGGAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGCTGACACTGTCGATGAAGCCCTGGAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACCAACAACAATAACAATAACAACAACCTCGGGATCGAGGGAAGGATTTCAAAATTC 2700
Db 2641 AACCAACAACAATAACAATAACAACAACCTCGGGATCGAGGGAAGGATTTCAAAATTC 2700

RESULT 10
ADO23603
ID ADO23603 standard; DNA; 7352 BP.
XX ADO23603;
XX 01-JUL-2004 (first entry)
XX DNA encoding MBP-ToxoP30del14C(52-294aa) fusion protein.
XX P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; gene.
XX Toxoplasma gondii.
OS Synthetic.
XX US2004067239-A1.
XX 08-APR-2004.
XX 02-OCT-2002; 2002US-00263153.
XX 02-OCT-2002; 2002US-00263153.
XX (MAIN/) MAINE G T.
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX WPI; 2004-304563/28.
DR P-PSDB; ADO23604.
XX Novel purified polypeptide having sequence identity to amino acid sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C, Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG antibodies to Toxoplasma gondii.
XX Example 2; Fig 19; 114pp; English.
XX The invention relates to a purified P30 antigen (I) chosen from 3 fully defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6 amino acids added to the C-terminus of the amino acid sequence of Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence of the amino acid sequence of Toxo30del13C P30 antigen sequence is substituted with alanine, or comprising the amino acid sequence chosen from MBP-Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1 fusion proteins. (I) is useful for detecting the presence of IgM antibodies to Toxoplasma gondii in a test sample, which involves contacting the test sample suspected of containing the IgM antibodies

CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
XX Toxop30 fusion protein of the invention.

SQ Sequence 7352 BP; 1896 A; 1852 C; 1943 G; 1661 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7352;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY	1	CGGACCACTCAATGGTGCAGAAACCTTTGCGGGTATGCGATGATGCGCCGGAAGAGA	60
DB	1	CGGACCACTCAATGGTGCAGAAACCTTTGCGGGTATGCGATGATGCGCCGGAAGAGA	60
QY	61	GTCAATTACAGGTGGTGAATGTGAACCAAGTAACTGATACAGATGTCGACAGTATGCG	120
DB	61	GTCAATTACAGGTGGTGAATGTGAACCAAGTAACTGATACAGATGTCGACAGTATGCG	120
QY	121	GTGTCCTTATCAGACCGGTTCCCGCGTGGTGAACCAAGCGAGCCACGTTTCTCGAAAA	180
DB	121	GTGTCCTTATCAGACCGGTTCCCGCGTGGTGAACCAAGCGAGCCACGTTTCTCGAAAA	180
QY	181	CGCGGAAAGTGAAGCGCGATGCGGAGCTGAATTACATCCCAACCGGTGGCAC	240
DB	181	CGCGGAAAGTGAAGCGCGATGCGGAGCTGAATTACATCCCAACCGGTGGCAC	240
QY	241	AACAACTGCGCGGCAACAGTCGTTGCTGATTTGGGTTGCCACTTCCAGTCTGCGCCCTGC	300
DB	241	AACAACTGCGCGGCAACAGTCGTTGCTGATTTGGGTTGCCACTTCCAGTCTGCGCCCTGC	300
QY	301	ACGCGCGTCGCAATTTGTCGCGCGATTAATCTCGCGCGATCAACTGCGGTGCCAGC	360
DB	301	ACGCGCGTCGCAATTTGTCGCGCGATTAATCTCGCGCGATCAACTGCGGTGCCAGC	360
QY	361	TGTTGCTGATGTAAGCAAGCGCGCTGAAAGCTGTAAAGCGCGGTGCACATC	420
DB	361	TGTTGCTGATGTAAGCAAGCGCGCTGAAAGCTGTAAAGCGCGGTGCACATC	420
QY	421	TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACCAAGATGCCA	480
DB	421	TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACCAAGATGCCA	480
QY	481	TTGCTGTGGAAGCTCGCTGCACTAATGTTCCGGGTTATTTCTTGATGTCCTGACACAGA	540
DB	481	TTGCTGTGGAAGCTCGCTGCACTAATGTTCCGGGTTATTTCTTGATGTCCTGACACAGA	540
QY	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGCACTGGCGGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGCACTGGCGGTGGAGCATC	600
QY	601	TGGTCGCAATTGGGTACACAGCAATTCGGCTGTTAGCGGGCCCAATTAAGTTCGTCTCGG	660
DB	601	TGGTCGCAATTGGGTACACAGCAATTCGGCTGTTAGCGGGCCCAATTAAGTTCGTCTCGG	660
QY	661	CGCGTCTGCGTCTGGCTGGGATTAATATCTCACTCCCAATCAATTTACGCCGATAG	720
DB	661	CGCGTCTGCGTCTGGCTGGGATTAATATCTCACTCCCAATCAATTTACGCCGATAG	720
QY	721	CGGAACCGGAAGCGCATGGAAGTCCCATGTCGGTTCCTCAACCAACCATGCAATGCTGA	780
DB	721	CGGAACCGGAAGCGCATGGAAGTCCCATGTCGGTTCCTCAACCAACCATGCAATGCTGA	780
QY	781	ATGAGGCACTGTTCCCACTGCGATGCTGGTTCGCAACGATCAGATGGCGCTGGCGCAA	840
DB	781	ATGAGGCACTGTTCCCACTGCGATGCTGGTTCGCAACGATCAGATGGCGCTGGCGCAA	840
QY	841	TGCGCGCATTAACGAGTCCGGCTGCGGCTGGTGGGATATCTCGGTAGTGGGATACG	900
DB	841	TGCGCGCATTAACGAGTCCGGCTGCGGCTGGTGGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACGGAAGACAGTCTATGTTATATCCGCGGTTAACCAATCAACAGGATTTTC	960
DB	901	ACGATACGGAAGACAGTCTATGTTATATCCGCGGTTAACCAATCAACAGGATTTTC	960

QY	961	GCCTGCTGGGGCAAAACAGGGTGACCCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
DB	961	GCCTGCTGGGGCAAAACAGGGTGACCCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
QY	1021	AGGCAATACAGCTGTTGCCGCTCTCACTGCTGTAAGAAAAAACCCCTTGGCGCCCAATA	1080
DB	1021	AGGCAATACAGCTGTTGCCGCTCTCACTGCTGTAAGAAAAAACCCCTTGGCGCCCAATA	1080
QY	1081	CGCAAAACCGCTCTCTCCCGCGGCTTGCGCGATTCATTAATGACGCTGGCAGCAGGTTT	1140
DB	1081	CGCAAAACCGCTCTCTCCCGCGGCTTGCGCGATTCATTAATGACGCTGGCAGCAGGTTT	1140
QY	1141	CCGCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTCAATTAG	1200
DB	1141	CCGCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTCAATTAG	1200
QY	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACCGTGCACCAATGCTTCTGGCG	1260
DB	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACCGTGCACCAATGCTTCTGGCG	1260
QY	1261	TCAGGACGCCATCGGAAGCTGTGATGCTGTGAGTGTGCGAGTGTAAATCACTGCAATAATCG	1320
DB	1261	TCAGGACGCCATCGGAAGCTGTGATGCTGTGAGTGTGCGAGTGTAAATCACTGCAATAATCG	1320
QY	1321	TGTCGCTCAAGGCGCACTCCCGTCTGGATAATGTTTTTGGCGCGGACATCAAAACGGTT	1380
DB	1321	TGTCGCTCAAGGCGCACTCCCGTCTGGATAATGTTTTTGGCGCGGACATCAAAACGGTT	1380
QY	1381	CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGTGGA	1440
DB	1381	CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGTGGA	1440
QY	1441	ATTGTGAGCGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTGGTGTTCACGA	1500
DB	1441	ATTGTGAGCGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTGGTGTTCACGA	1500
QY	1501	GCATTTACCAACCAAGGACCATAGATATGAAATCTGAAAGCTGAAAGGTTAACTGGTAACTCG	1560
DB	1501	GCATTTACCAACCAAGGACCATAGATATGAAATCTGAAAGCTGAAAGGTTAACTGGTAACTCG	1560
QY	1561	ATTAAACCGGATTAAGGCTATTAACCGTCTCGCTGAAAGTCCGTTAAGAAATTCGAGAAAGAT	1620
DB	1561	ATTAAACCGGATTAAGGCTATTAACCGTCTCGCTGAAAGTCCGTTAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT	1680
DB	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT	1680
QY	1681	CGGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCAACAGCCGCTTTGGTGGCTAC	1740
DB	1681	CGGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCAACAGCCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCTGTTGGCTGAAATCAACCGGCAAAAGCGTTCCAGGACAGCTGTAT	1800
DB	1741	GCTCAATCTGGCTGTTGGCTGAAATCAACCGGCAAAAGCGTTCCAGGACAGCTGTAT	1800
QY	1801	CGTTTTACCTGGGATGCGGTAGCTTACACGCGCAAGCTGATTTGCTTACCGATCGCTGTT	1860
DB	1801	CGTTTTACCTGGGATGCGGTAGCTTACACGCGCAAGCTGATTTGCTTACCGATCGCTGTT	1860
QY	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA	1920
DB	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA	1920
QY	1921	GAGATCCCGCGCTGGATAAGAACTGAAGGAAAGGTAAGAGCGCGCTGATGTTCAAC	1980
DB	1921	GAGATCCCGCGCTGGATAAGAACTGAAGGAAAGGTAAGAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTGGCGGCTGATTTGCTGCTGACGGGGTTATCGGTTCAAG	2040
DB	1981	CTGCAAGAACCGTACTTCACTGGCGGCTGATTTGCTGCTGACGGGGTTATCGGTTCAAG	2040

Qy	2041	TATGAAACGGCAAGTACGACATTAAGACACGTGGCGTGGATACGCTGGCGCGAAGCG	2100
Db	2041	TATGAAACGGCAAGTACGACATTAAGACACGTGGCGTGGATACGCTGGCGCGAAGCG	2100
Qy	2101	GGTCTGACCTTCCTGGTGTGACCTGATTAAAAACAAACACATGAATGCAGACACCGATTAC	2160
Db	2101	GGTCTGACCTTCCTGGTGTGACCTGATTAAAAACAAACACATGAATGCAGACACCGATTAC	2160
Qy	2161	TCCATCGCAGAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG	2220
Db	2161	TCCATCGCAGAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG	2220
Qy	2221	GCATGGTCCACATCGACACGACGAAAGTGAATTATGGTTAAACGGTACTCGACCTTC	2280
Db	2221	GCATGGTCCACATCGACACGACGAAAGTGAATTATGGTTAAACGGTACTCGACCTTC	2280
Qy	2281	AAGGTCACACCATCAAAACCGTTCTTGGCGTGTGAGCGCAGGTATTAAACGCCGCCAGT	2340
Db	2281	AAGGTCACACCATCAAAACCGTTCTTGGCGTGTGAGCGCAGGTATTAAACGCCGCCAGT	2340
Qy	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG	2400
Db	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG	2400
Qy	2401	GAAGCGTTAATAAGACAAACCGCTGGCTGCGTGTGAGCTGAACTCTTACGAGAAAG	2460
Db	2401	GAAGCGTTAATAAGACAAACCGCTGGCTGCGTGTGAGCTGAACTCTTACGAGAAAG	2460
Qy	2461	TTGGCGAAAGATCCACGCTATTGCGCCACCATGGAACGCCAGAAAGGTGAATCATG	2520
Db	2461	TTGGCGAAAGATCCACGCTATTGCGCCACCATGGAACGCCAGAAAGGTGAATCATG	2520
Qy	2521	CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGTGTGCTGCTGCTGCTGCTGCTG	2580
Db	2521	CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGTGTGCTGCTGCTGCTGCTGCTG	2580
Qy	2581	CCAGCGCTGCTGACACTGTGATGAAGCCCTGAAAGAGCGCAGACTAATTCGAGCTCG	2640
Db	2581	CCAGCGCTGCTGACACTGTGATGAAGCCCTGAAAGAGCGCAGACTAATTCGAGCTCG	2640
Qy	2641	AACAAACAACAATAACAATAACAATAACAATAACAATAACAATAACAATAACAATA	2700
Db	2641	AACAAACAACAATAACAATAACAATAACAATAACAATAACAATAACAATAACAATA	2700

RESULT 11

ADO23598

ID ADO23598 standard; DNA; 7370 BP.

AC ADO23598;

XX ADO23598;

XX ADO23598;

DT 01-JUL-2004 (first entry)

DE DNA encoding MBP-ToxoP30del13C(52-300aa) fusion protein.

XX P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;

XX MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;

KW Toxoplasma gondii; db; gene.

XX Toxoplasma gondii.

OS Synthetic.

XX US2004067239-A1.

XX 08-APR-2004.

XX 02-OCT-2002; 2002US-00263153.

XX 02-OCT-2002; 2002US-00263153.

XX (MAIN/) MAINE G T.

PA (PATE/) PATEL C B.

PA (GINS/) GINSBURG S R.

PA	(BLIE/) BLIESE T R.
XX	Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX	WPI; 2004-304563/28.
DR	P-PSDB; ADO23599.
XX	Novel purified polypeptide having sequence identity to amino acid
PT	sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
PT	Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
PT	antibodies to Toxoplasma gondii.
XX	Example 2; Fig 15; 114pp; English.
PS	The invention relates to a purified P30 antigen (I) chosen from 3 fully
CC	defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
CC	amino acids added to the C-terminus of the amino acid sequence of
CC	Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
CC	which at least one of the five C-terminal cysteine amino acids of the
CC	amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
CC	with alanine, or comprising the amino acid sequence chosen from MBP-
CC	Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
CC	fusion proteins. (I) is useful for detecting the presence of IgM
CC	antibodies to Toxoplasma gondii in a test sample, which involves
CC	contacting the test sample suspected of containing the IgM antibodies
CC	with a composition comprising (I) and detecting the presence of (I)/IgM
CC	antibody complexes. The present sequence represents DNA encoding a MBP-
CC	ToxoP30 fusion protein of the invention.
XX	Sequence 7370 BP; 1901 A; 1855 C; 1949 G; 1665 T; 0 U; 0 Other;
SQ	
Query Match 81.5%; Score 2690.4; DB 12; Length 7370;	
Best Local Similarity 99.8%; Pred. No. 0;	
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;	
Qy	1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGTATGGCATGATAGCCCGGAAGAGA 60
Db	1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGTATGGCATGATAGCCCGGAAGAGA 60
Qy	61 GTCAATTCCAGGTGGTGAATGTGAAACCCAGTAAAGCTTATAGCATGTCGACAGTATGCCG 120
Db	61 GTCAATTCCAGGTGGTGAATGTGAAACCCAGTAAAGCTTATAGCATGTCGACAGTATGCCG 120
Qy	121 GTGTCCTTATCAGACCGTTTCCCGCGTGTGAAACCCAGTAAAGCTTATAGCATGTCGACAGTATGCCG 180
Db	121 GTGTCCTTATCAGACCGTTTCCCGCGTGTGAAACCCAGTAAAGCTTATAGCATGTCGACAGTATGCCG 180
Qy	181 CGCGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC 240
Db	181 CGCGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC 240
Qy	241 AACAACTGGCGGCAACACAGTCGTTGCTGATTTGGGCTTGGCCACTTCCAGTCTGGCCCTGCG 300
Db	241 AACAACTGGCGGCAACACAGTCGTTGCTGATTTGGGCTTGGCCACTTCCAGTCTGGCCCTGCG 300
Qy	301 ACGCCCGCTCGCAAAATTGTTCGCGCGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
Db	301 ACGCCCGCTCGCAAAATTGTTCGCGCGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
Qy	361 TGGTGGTGTGATGGTAGAACGAGCGCGTGAAGCCCTGTAAAGCGCGGTGCAACAATC 420
Db	361 TGGTGGTGTGATGGTAGAACGAGCGCGTGAAGCCCTGTAAAGCGCGGTGCAACAATC 420
Qy	421 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATTAATCTCGCTGATGATCAACAGGATGCCA 480
Db	421 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATTAATCTCGCTGATGATCAACAGGATGCCA 480
Qy	481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTATTCTTGAATGCTCTGACACAGA 540
Db	481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTATTCTTGAATGCTCTGACACAGA 540
Qy	541 CACCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACCGGACTGGCGGTGGAGCATC 600
PA	

Db 541 CACCCATCAACAGTATTATTTCTCCATGAAGACGGTACGGGCTGGAGCATC 600
Qy 601 TGGTCGCAATTGGGTACACGAAATCGCGCTGTAGCGGCCCATTAAGTTCTGCTCGG 660
Db 601 TGGTCGCAATTGGGTACACGAAATCGCGCTGTAGCGGCCCATTAAGTTCTGCTCGG 660
Qy 661 CGGCTCTGGTCTGGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGGCTCTGGTCTGGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CCGAAGCGGAAGCGGACGTGGAGTGCATGTCGGTTCCTCAACAAACCATGCAATGCTGA 780
Db 721 CCGAAGCGGAAGCGGACGTGGAGTGCATGTCGGTTCCTCAACAAACCATGCAATGCTGA 780
Qy 781 ATGAGGGCATCTTCCCACTGCGATGCTGGTTGCCAAGATCAGATGCGCTGGCGCGAA 840
Db 781 ATGAGGGCATCTTCCCACTGCGATGCTGGTTGCCAAGATCAGATGCGCTGGCGCGAA 840
Qy 841 TCGCGGCCATTACCGAGTCCGGCTGGCGGTGGTGGCGGATATCTCGGTAGTGGGATACG 900
Db 841 TCGCGGCCATTACCGAGTCCGGCTGGCGGTGGTGGCGGATATCTCGGTAGTGGGATACG 900
Qy 901 AGCATACCGAAGACAGCTCATGTATTATCCCGCGCTTAACACCATCAAAAGGATTTTC 960
Db 901 AGCATACCGAAGACAGCTCATGTATTATCCCGCGCTTAACACCATCAAAAGGATTTTC 960
Qy 961 GCCTGCTGGGGAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGCTGGGGAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGCGCTCTCACTGGTGAAGAAACCAACCTGCGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGCGCTCTCACTGGTGAAGAAACCAACCTGCGCGCCCAATA 1080
Qy 1081 CGCAACCGCCTCTCCCGCGCTTGGCGGATTAATTAATGAGTGGCAGCAGAGTTT 1140
Db 1081 CGCAACCGCCTCTCCCGCGCTTGGCGGATTAATTAATGAGTGGCAGCAGAGTTT 1140
Qy 1141 CCGACTGGAAGCGGCGAGTGGGGAACCGCAATTAATGAGTGGTACCTCACTATTAG 1200
Db 1141 CCGACTGGAAGCGGCGAGTGGGGAACCGCAATTAATGAGTGGTACCTCACTATTAG 1200
Qy 1201 GCACAATCTCATGTTTGACAGCTTATCATGCTGACGCTGACCAATGCTTCTGCGG 1260
Db 1201 GCACAATCTCATGTTTGACAGCTTATCATGCTGACGCTGACCAATGCTTCTGCGG 1260
Qy 1261 TCAGGAGCCATCGGAAGCTGTGATGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGATGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1320
Qy 1321 TGTGCTCAAGCGCCTCTCCCGTCTCGATATGTTTTCGCGCGACATCAATCGGTT 1380
Db 1321 TGTGCTCAAGCGCCTCTCCCGTCTCGATATGTTTTCGCGCGACATCAATCGGTT 1380
Qy 1381 CTGCAAAATTTCTGAAATGAGCTGTGCAAAATTAATCATCGGCTCGTAAATGTTGGA 1440
Db 1381 CTGCAAAATTTCTGAAATGAGCTGTGCAAAATTAATCATCGGCTCGTAAATGTTGGA 1440
Qy 1441 ATTGTAGCGGATCAAAATTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Db 1441 ATTGTAGCGGATCAAAATTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Qy 1501 GCATCTCACCAAGAGCATGATGATGAAACTGAAGAAGTAACTGGTAAATCTCG 1560
Db 1501 GCATCTCACCAAGAGCATGATGATGAAACTGAAGAAGTAACTGGTAAATCTCG 1560
Qy 1561 ATTAACGGCGATTAAGGCTATAACGGTCTCGCTGAGTTCGCTGAGTTCGAGTTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATTAAGGCTATAACGGTCTCGCTGAGTTCGCTGAGTTCGAGTTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGCTACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGCTACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCACAGGTT 1680

RESULT 12

AD023639

ID AD023639 standard; DNA; 7370 BP.

Qy 1681 CCGGAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGTAC 1740
Db 1681 CCGGAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGTAC 1740
Qy 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCGCAAAAGGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCGCAAAAGGTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCCGTACGTTTCAACGGCAAGCTGATTTGCTTACCAGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCCGTACGTTTCAACGGCAAGCTGATTTGCTTACCAGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGCTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGCTAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGCTGACGGGGTTATGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGCTGACGGGGTTATGGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTAAGAGCTGGCGTGGATAAACGCTGGCGGAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGAGCTGGCGTGGATAAACGCTGGCGGAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAACAAACACATGATGACAGACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAACAAACACATGATGACAGACCGGATTAC 2160
Qy 2161 TCCATGCGAGAACTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCCCGCTGG 2220
Db 2161 TCCATGCGAGAACTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCCCGCTGG 2220
Qy 2221 GCATGCTCAACATCGACACCAGCAAAAGTGAATTATGGTGAACGGTACTGCGGACTTC 2280
Db 2221 GCATGCTCAACATCGACACCAGCAAAAGTGAATTATGGTGAACGGTACTGCGGACTTC 2280
Qy 2281 AAGGGTCAACCTATCCAAACCGTTGCTGGCTGCTGAGCGCAGGTATTAAACGCCCGCAGT 2340
Db 2281 AAGGGTCAACCTATCCAAACCGTTGCTGGCTGCTGAGCGCAGGTATTAAACGCCCGCAGT 2340
Qy 2341 CCGAACAAGAGCTGCAAAAGAGTTCTCGAATACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAGAGCTGCAAAAGAGTTCTCGAATACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGTTTAAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGTTTAAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGGGCAAGATCCAGGTATTGCGCCACCATGGAACCGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGGCAAGATCCAGGTATTGCGCCACCATGGAACCGCCAGAAAGGTGAAATCATG 2520
Qy 2521 CCGAATCCCGCAGATGCTCCGCTTCTGCTGATGCGTGGCTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAATCCCGCAGATGCTCCGCTTCTGCTGATGCGTGGCTACTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACACTGTGCGTGAAGCCCTGAAAGAGCGCGACAGTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACACTGTGCGTGAAGCCCTGAAAGAGCGCGACAGTAAATTCGAGCTCG 2640
Qy 2641 AACAACAACAACAATAAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 2700
Db 2641 AACAACAACAACAATAAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 2700

XX	ADO23639;
AC	
XX	
XX	01-JUL-2004 (first entry)
DT	
XX	
DE	DNA encoding MBP-Toxop30MIX1 fusion protein.
XX	
KW	P30 antigen; Toxo30del13C; Toxo30del12C; Toxop30 MIX1;
KW	MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW	Toxoplasma gondii; ds; gene.
XX	
OS	Toxoplasma gondii.
OS	Synthetic.
XX	
PN	US2004067239-A1.
XX	
PD	08-APR-2004.
XX	
PF	02-OCT-2002; 2002US-00263153.
XX	
PR	02-OCT-2002; 2002US-00263153.
XX	
PA	(MAIN/) MAINE G T.
PA	(PATE/) PATEL C B.
PA	(GINS/) GINSBURG S R.
PA	(BLIE/) BLIESE T R.
XX	
PI	Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX	
DR	WPI; 2004-304563/28.
DR	P-PSDB; ADO23640.
XX	
PT	Novel purified polypeptide having sequence identity to amino acid
PT	sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
PT	Toxo30del12C, Toxop30 MIX1, useful for detecting presence of IgM or IgG
PT	antibodies to Toxoplasma gondii.
XX	
XX	Example 5; Fig 31; 114pp; English.
XX	
CC	The invention relates to a purified p30 antigen (I) chosen from 3 fully
CC	defined Toxo30del13C, Toxo30del12C and Toxop30 MIX1 sequences, having 1-6
CC	amino acids added to the C-terminus of the amino acid sequence of
CC	Toxo30del12C p30 antigen sequence, or comprising an amino acid sequence in
CC	which at least one of the five C-terminal cysteine amino acids of the
CC	amino acid sequence of Toxo30del13C p30 antigen sequence is substituted
CC	with alanine, or comprising the amino acid sequence chosen from MBP-
CC	Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
CC	fusion proteins. (I) is useful for detecting the presence of IgM
CC	antibodies to Toxoplasma gondii in a test sample, which involves
CC	contacting the test sample suspected of containing the IgM antibodies
CC	with a composition comprising (I) and detecting the presence of (I)/IgM
CC	antibody complexes. The present sequence represents DNA encoding a MBP-
CC	Toxop30 fusion protein of the invention.
XX	
SQ	Sequence 7370 BP; 1900 A; 1856 C; 1950 G; 1664 T; 0 U; 0 Other;
	Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
	Best Local Similarity 99.8%; Pred No. 0;
	Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0
Qy	1 CCGACACCATCGAATTGGTGTCAAAACCTTTTCGGGTATGCAGTATGCGCCCGGAAGA 60
Db	1 CGGACACCATCGAATTGGTGTCAAAACCTTTTCGGGTATGCAGTATGCGCCCGGAAGA 60
Qy	61 GTCAATTCAGGTGTGAATGTGAAACCAAGTAACGTTATACGATGTCGAGAGTATGCCG 120
Db	61 GTCAATTCAGGTGTGAATGTGAAACCAAGTAACGTTATACGATGTCGAGAGTATGCCG 120
Qy	121 GTGTCTCTTATCAGACCGTTTTCCCGCGTGTTGAACCGCCAGCACCGTTTCTGCGAAAA 180
Db	121 GTGTCTCTTATCAGACCGTTTTCCCGCGTGTTGAACCGCCAGCACCGTTTCTGCGAAAA 180
Qy	181 CGCGGAAAAAGTGGAAACGCGCATGTGGCGGAGCTGAATTTACATTTCCCACCGCGTGGCAC 240

CC antibodies to *Toxoplasma gondii* in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (i) and detecting the presence of (i)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC Toxop30 fusion protein of the invention.
XX
SQ Sequence 7370 BP; 1900 A; 1857 C; 1950 G; 1663 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCGACACCATGAAATGGTGCAAAACCTTTCCGGGTATGCGATGATAGCGCCCGGAAGAGA 60
DB 1 CCGACACCATGAAATGGTGCAAAACCTTTCCGGGTATGCGATGATAGCGCCCGGAAGAGA 60

QY 61 GTCAATTCAGGGTGTGAATGTGAACACCAAGTAACGTTATACGATGTCGACAGTATGCGG 120
DB 61 GTCAATTCAGGGTGTGAATGTGAACACCAAGTAACGTTATACGATGTCGACAGTATGCGG 120

QY 121 GTGTCCTTATCAGACCGTTTCCCGGGTGTGAACACCGCCAGCCACGCTTTCTCGAAAA 180
DB 121 GTGTCCTTATCAGACCGTTTCCCGGGTGTGAACACCGCCAGCCACGCTTTCTCGAAAA 180

QY 181 CGCGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTTACATTTCCCAACCGCTGGGCAC 240
DB 181 CGCGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTTACATTTCCCAACCGCTGGGCAC 240

QY 241 AACAACTGCGGGCAAAACAGTCTGTGCTGATTTGGGGTGGCCACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGCGGGCAAAACAGTCTGTGCTGATTTGGGGTGGCCACCTCCAGTCTGGCCCTGC 300

QY 301 AGCGCCGTCGCAAAATGTCGCGGCGATTAATCTCGCGCGGATCAATCGGTCGACGCG 360
DB 301 AGCGCCGTCGCAAAATGTCGCGGCGATTAATCTCGCGCGGATCAATCGGTCGACGCG 360

QY 361 TGGTGGTGTGATGCTAGAACGAAGCGGCTCGAAGCCGTGAAGCGGCGTGCACATC 420
DB 361 TGGTGGTGTGATGCTAGAACGAAGCGGCTCGAAGCCGTGAAGCGGCGTGCACATC 420

QY 421 TTCTCGCGCAACGCTGAGTGGGCTGATCATTAATCTATCCGCTGGATGACCAAGTACCA 480
DB 421 TTCTCGCGCAACGCTGAGTGGGCTGATCATTAATCTATCCGCTGGATGACCAAGTACCA 480

QY 481 TTGCTGTGGAAGCTGCTGCAATAATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA 540
DB 481 TTGCTGTGGAAGCTGCTGCAATAATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA 540

QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGGCACTGGGCGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGGCACTGGGCGTGGAGCATC 600

QY 601 TGGTCGCAATGGGTGTCACAGCAAAATCGCGCTGTAGCGGCGCCATTTAAGTCTGTCTCGG 660
DB 601 TGGTCGCAATGGGTGTCACAGCAAAATCGCGCTGTAGCGGCGCCATTTAAGTCTGTCTCGG 660

QY 661 CGCGTCTGCGTCTGGCTGGCTGGCAATAATCTCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGCGTCTGGCTGGCTGGCAATAATCTCTCACTCGCAATCAAAATTCAGCCGATAG 720

QY 721 CGGAACGGGAAGCGGACTGGAGTGCCATGTCGCGTGTGGTGGGATATCTCGTAGTGGGATCG 780
DB 721 CGGAACGGGAAGCGGACTGGAGTGCCATGTCGCGTGTGGTGGGATATCTCGTAGTGGGATCG 780

QY 781 ATGAGGGCATCGTTCCCACTGGCATGCTGGTGGCCAAACGATCAGATGGGCGCTGGGCGCAA 840
DB 781 ATGAGGGCATCGTTCCCACTGGCATGCTGGTGGCCAAACGATCAGATGGGCGCTGGGCGCAA 840

QY 841 TGGCGGCCATTACCGAGTCCGGGCTCGCGGTGGTGGGATATCTCGTAGTGGGATCG 900
DB 841 TGGCGGCCATTACCGAGTCCGGGCTCGCGGTGGTGGGATATCTCGTAGTGGGATCG 900

QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCAACCATCAAAACAGGATTTTC 960

DB 901 ACGATACCGAAGACAGCTCATGTTATATATCCCGCGGTTAAACCAACCATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGGCAAAACGAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGGTGA 1020
DB 961 GCCTGCTGGGGCAAAACGAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGGTGA 1020

QY 1021 AGGCAATCAGCTGTTGCCCGCTCTCACTGTGTGAAAGAAAAACCAACCTCTGGCGCCCAATA 1080
DB 1021 AGGCAATCAGCTGTTGCCCGCTCTCACTGTGTGAAAGAAAAACCAACCTCTGGCGCCCAATA 1080

QY 1081 CGCAAAACCGCCTCTCCCCCGCGCTTGGCCGATTCATTAATGCAGCTGGCAGCAGAGTTT 1140
DB 1081 CGCAAAACCGCCTCTCCCCCGCGCTTGGCCGATTCATTAATGCAGCTGGCAGCAGAGTTT 1140

QY 1141 CCCGACTGGAAGGGGCGAGTGAGGCGAACCGCAATTAATGTGAGTTAGTACTACTCATTTAG 1200
DB 1141 CCCGACTGGAAGGGGCGAGTGAGGCGAACCGCAATTAATGTGAGTTAGTACTACTCATTTAG 1200

QY 1201 GCACAAATCTCATGTTTTCAGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260
DB 1201 GCACAAATCTCATGTTTTCAGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260

QY 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGCAGTTCGTAAATCACTGTCATTAATTCG 1320
DB 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGCAGTTCGTAAATCACTGTCATTAATTCG 1320

QY 1321 TGTGCTCAAGCGGCACCTCCCGTCTTGATTAATGTTTTTGGCGCGACATCATNAACGGTT 1380
DB 1321 TGTGCTCAAGCGGCACCTCCCGTCTTGATTAATGTTTTTGGCGCGACATCATNAACGGTT 1380

QY 1381 CTGGCAATATCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGTCGA 1440
DB 1381 CTGGCAATATCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGTCGA 1440

QY 1441 ATTGTGACGGATTAACAAATTTTCACACAGAAACAGCCAGTCCGTTTAGTGTGTTTTCACGA 1500
DB 1441 ATTGTGACGGATTAACAAATTTTCACACAGAAACAGCCAGTCCGTTTAGTGTGTTTTCACGA 1500

QY 1501 GCATCTTCCAAACAAAGGACCATAGATTATGAAAACTGAAGAAAGGTAAACTGTGTAATCTGG 1560
DB 1501 GCATCTTCCAAACAAAGGACCATAGATTATGAAAACTGAAGAAAGGTAAACTGTGTAATCTGG 1560

QY 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAGTGGTAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAGTGGTAAGAAATTCGAGAAAGAT 1620

QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCGCCACAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCGCCACAGGTT 1680

QY 1681 GCGGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACGACCGCTTTGGTGGCTAC 1740
DB 1681 GCGGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACGACCGCTTTGGTGGCTAC 1740

QY 1741 GCTCAATCTGGCGCTGTGGCTGAAAATCAACCCCGGCAAAAGCGTTCCAGACAAGCTGTAT 1800
DB 1741 GCTCAATCTGGCGCTGTGGCTGAAAATCAACCCCGGCAAAAGCGTTCCAGACAAGCTGTAT 1800

QY 1801 CCGTTTACCTGGGATGCGGTACGTTTACACGCGAAGCTGATTGCTTACCCGATCGCTGTT 1860
DB 1801 CCGTTTACCTGGGATGCGGTACGTTTACACGCGAAGCTGATTGCTTACCCGATCGCTGTT 1860

QY 1861 GAAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAACCCCGCAAAAAACCTGGGAA 1920
DB 1861 GAAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAACCCCGCAAAAAACCTGGGAA 1920

QY 1921 GAGATCCCGCGCTCGGATAAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
DB 1921 GAGATCCCGCGCTCGGATAAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980

QY 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATTGCTGTCGACGGGGTGTATCGGTTCAAG 2040

```
Db 1981 CTGAAGAACGTAATTACCTGGCGCTGATTGCTGCTGAGCGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAACGCAAGTACGACATTAAAGACGCTGGCGTGGATACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACGCAAGTACGACATTAAAGACGCTGGCGTGGATACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAACAAACACATCAATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAACAAACACATCAATGACAGACACCGATTAC 2160
Qy 2161 TCCATCGAAGAGCTGCTTTAATAAGCGGAAACAGCGATGACATCAACAGCGCCGTGG 2220
Db 2161 TCCATCGAAGAGCTGCTTTAATAAGCGGAAACAGCGATGACATCAACAGCGCCGTGG 2220
Qy 2221 GCATGGTCAACATCGACACAGCAAGTGAATTATGGTGAACGGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACAGCAAGTGAATTATGGTGAACGGTACTGCCGACCTTC 2280
Qy 2281 AAGGGTCAACCATCAAAACCGTTCTGGTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGGTCAACCATCAAAACCGTTCTGGTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGCGGAAAGATCCACGTATTGCGGCCACCATGGAAGAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGCGGAAAGATCCACGTATTGCGGCCACCATGGAAGAAAGGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCGCGAGATGTCGCTTTCTGGTATCCGCTAGCTGCGGTGATCAAGCC 2580
Db 2521 CCGAACATCCGCGAGATGTCGCTTTCTGGTATCCGCTAGCTGCGGTGATCAAGCC 2580
Qy 2581 GCCAGCGGTGTCAGACTGTGATGAAGCCCTGAAAGACGGCAGACTAATTGAGCTCG 2640
Db 2581 GCCAGCGGTGTCAGACTGTGATGAAGCCCTGAAAGACGGCAGACTAATTGAGCTCG 2640
Qy 2641 AACAAACAAACAAATACATTAACAACAACCTCGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAAACAAATACATTAACAACAACCTCGGATCGAGGGAAGGATTCAGAAATTC 2700
```

RESULT 14
ADO23644

ID ADO23644 standard; DNA; 7370 BP.

XX AC ADO23644;

XX DT 01-JUL-2004 (first entry)

XX DE DNA encoding MBP-Toxop30MIX3 fusion protein.

XX KW P30 antigen; Toxo30del13C; Toxo30del12C; Toxop30 MIX1;
XX KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
XX KW Toxoplasma gondii; da; gene.

XX OS Toxoplasma gondii.

XX PN Synthetic.

XX PN US2004067239-A1.

XX XX 08-APR-2004.

XX XX 02-OCT-2002; 2002US-00263153.

XX XX 02-OCT-2002; 2002US-00263153.

XX PR (MAIN/) MAINE G T.

PA

(PATE/) PATEL C B.
(GINS/) GINSBURG S R.
(BLIE/) BLIESE T R.

Maine GT, Patel CB, Ginsburg SR, Bliese TR;

WPI; 2004-304563/28.

P-PSDB; ADO23645.

Novel purified polypeptide having sequence identity to amino acid
sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
Toxo30del12C, Toxop30 MIX1, useful for detecting presence of IgM or IgG
antibodies to Toxoplasma gondii.

Example 5; Fig 33; 114pp; English.

The invention relates to a purified P30 antigen (I) chosen from 3 fully
defined Toxo30del13C, Toxo30del12C and Toxop30 MIX1 sequences, having 1-6
amino acids added to the C-terminus of the amino acid sequence of
Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
which at least one of the five C-terminal cysteine amino acids of the
amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
with alanine, or comprising the amino acid sequence chosen from MBP-
Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
fusion proteins. (I) is useful for detecting the presence of IgM
antibodies to Toxoplasma gondii in a test sample, which involves
contacting the test sample suspected of containing the IgM antibodies
with a composition comprising (I) and detecting the presence of (I)/IgM
antibody complexes. The present sequence represents DNA encoding a MBP-
Toxop30 fusion protein of the invention.

SQ Sequence 7370 BP; 1900 A; 1857 C; 1950 G; 1663 T; 0 U; 0 Other;

Query Match

Best Local Similarity 81.5%; Score 2690.4; DB 12; Length 7370;

Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy 1 CCACACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATATAGCGCCCGAAGAGA 60
Db 1 CCACACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATATAGCGCCCGAAGAGA 60
Qy 61 GTCAATTCAGGGTGGTGAATGTGAACCCAGTATACGATGTGCGAGAGTATGCCG 120
Db 61 GTCAATTCAGGGTGGTGAATGTGAACCCAGTATACGATGTGCGAGAGTATGCCG 120
Qy 121 GTGTCTCTTATCAGACCGTTTCCGGGTGTGAACCCAGGCGACGCTTCTTCGCGAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCGGGTGTGAACCCAGGCGACGCTTCTTCGCGAAA 180
Qy 181 CGCGGGAAGGTTGAAGCGCGGATGCGGAGCTGAATTTACATTCCCAACCGCGTGGCAC 240
Db 181 CGCGGGAAGGTTGAAGCGCGGATGCGGAGCTGAATTTACATTCCCAACCGCGTGGCAC 240
Qy 241 AACAACTGGCGGCAAAACAGTCTGCTGATTGGCGTTGCGACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAAAACAGTCTGCTGATTGGCGTTGCGACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGCCGCTCGCAAAATGTCGGCGGATTAATCTCGCGCGGATCACTGGGTGCCAGCG 360
Db 301 ACGCCGCTCGCAAAATGTCGGCGGATTAATCTCGCGCGGATCACTGGGTGCCAGCG 360
Qy 361 TGGTGGTGTGATGGTAGAACGAAAGCGGCTGCAAGCCCTGTAAGCGCGGTCACAAATC 420
Db 361 TGGTGGTGTGATGGTAGAACGAAAGCGGCTGCAAGCCCTGTAAGCGCGGTCACAAATC 420
Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATTAATCTATCGGTGGATGACAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATTAATCTATCGGTGGATGACAGGATGCCA 480
Qy 481 TTGCTGTGAAGCTGCTGCACCTAATGTTCCGCGGTTATTTCTTGATGCTCTGACGAGA 540
Db 481 TTGCTGTGAAGCTGCTGCACCTAATGTTCCGCGGTTATTTCTTGATGCTCTGACGAGA 540
```


QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAAGCGGTACGGCACTGGCGCTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAAGCGGTACGGCACTGGCGCTGGAGCATC 600
QY 601 TGGTCGCATTTGGGTACACAGCAAAATCGGCTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Db 601 TGGTCGCATTTGGGTACACAGCAAAATCGGCTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGCTGTGGCTGGCGTGGCAATAAATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db 661 CGCGTCTCGCTGTGGCTGGCGTGGCAATAAATCTCACTCGCAATCAAAATTCAGCGGATAG 720
QY 721 CGGAACGGGAAGGCACTGAGGTGCGATTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGCACTGAGGTGCGATTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCTTCCCACTGCGATGCTGGTTGCCAAGATCAGATGGCGCTGGCGCGCA 840
Db 781 ATGAGGGCATCTTCCCACTGCGATGCTGGTTGCCAAGATCAGATGGCGCTGGCGCGCA 840
QY 841 TGGCGGCCATTACCGAGTCCGGGCTGGCGGTTGGTGCGGATATCTCGGTAGTGGGATAG 900
Db 841 TGGCGGCCATTACCGAGTCCGGGCTGGCGGTTGGTGCGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACATCAACAGGATTTTC 960
QY 961 GCCTCTGGGGCAACACGAGCTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTCTGGGGCAACACGAGCTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAGAAACCCCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAGAAACCCCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGCACTGGCAACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGCACTGGCAACAGGTTT 1140
QY 1141 CCCGACTGAAAGCGGCGAGTGCAGCGCAACGAATTAATGTGAGTTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGAAAGCGGCGAGTGCAGCGCAACGAATTAATGTGAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAATTCATGTTTGAACGCTTATCATGACTGCAGGTGCACCAATGCTCTGGCG 1260
Db 1201 GCACAATTCATGTTTGAACGCTTATCATGACTGCAGGTGCACCAATGCTCTGGCG 1260
QY 1261 TCAGGCAGGCATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAAATCACTGCATTAATCG 1320
Db 1261 TCAGGCAGGCATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAAATCACTGCATTAATCG 1320
QY 1321 TGTCTCTCAAGCGCACTCCCGTCTGGATATGTTTTTTCGCGCGACATCAATACGTT 1380
Db 1321 TGTCTCTCAAGCGCACTCCCGTCTGGATATGTTTTTTCGCGCGACATCAATACGTT 1380
QY 1381 CTGGCAAAATATTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
Db 1381 CTGGCAAAATATTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
QY 1441 ATTGTAGCGGATACAAATTCACACAGAAACAGCCAGTCCGTTAGGTGTTTTCAAGA 1500
Db 1441 ATTGTAGCGGATACAAATTCACACAGAAACAGCCAGTCCGTTAGGTGTTTTCAAGA 1500
QY 1501 GCACCTCCACCAAGGACCATAGCATATGAAAAATCGAAGAAAGGTAACCTGGTAATCTGG 1560
Db 1501 GCACCTCCACCAAGGACCATAGCATATGAAAAATCGAAGAAAGGTAACCTGGTAATCTGG 1560
QY 1561 ATTTAAGCGGATATAAGGCTTAAACCGTCTCGCTGAAAGTCCGTTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTTAAGCGGATATAAGGCTTAAACCGTCTCGCTGAAAGTCCGTTAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAAGAGAAATTCACAGGTT 1680

Db 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAAGAGAAATTCACAGGTT 1680
QY 1681 CGGCAACTGGCGATGGCCCTGACATATATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGGCGATGGCCCTGACATATATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGGCTAGCTTTACAAACGCAAGCTGATTTGCTTACCCGATCCCTGTT 1860
Db 1801 CCGTTTACCTGGGATGGCTAGCTTTACAAACGCAAGCTGATTTGCTTACCCGATCCCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATATAAGAACTGAAAGGAAAGGTAAAGCGCGCTGATTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATATAAGAACTGAAAGGAAAGGTAAAGCGCGCTGATTTCAAC 1980
QY 1981 CTGCAAGAACGCTACTTCACTGGCGCTGATTTGCTGCTGACGGGGGTTATCGCTTCAAG 2040
Db 1981 CTGCAAGAACGCTACTTCACTGGCGCTGATTTGCTGCTGACGGGGGTTATCGCTTCAAG 2040
QY 2041 TATGAAAAACGCAATGACATTAACACGCTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGCAATGACATTAACACGCTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAAACAAACATGAAATGCGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAAACAAACATGAAATGCGAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCCCTTTAATAAGGCGCAACAGCGATGACCATCAAGCGGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCCCTTTAATAAGGCGCAACAGCGATGACCATCAAGCGGCCGCTGG 2220
QY 2221 GCATGCTCAACATCGACACCAACGTAATTAATGCTGTAACCGGTACTGCGCACTTTC 2280
Db 2221 GCATGCTCAACATCGACACCAACGTAATTAATGCTGTAACCGGTACTGCGCACTTTC 2280
QY 2281 AAGGTCAAACATCAAAACCGTTCTGGCGTGTGAGCGCAGGTATTAAACGCCCACT 2340
Db 2281 AAGGTCAAACATCAAAACCGTTCTGGCGTGTGAGCGCAGGTATTAAACGCCCACT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGCTCG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGCTCG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCACGTTATTCGCGCCACCATGGAAGAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATTCGCGCCACCATGGAAGAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCGCGAGATGTCGCTTTCTGGTATGCGCGTACCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCGCGAGATGTCGCTTTCTGGTATGCGCGTACCTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACCAACAAACAAATAACAACTCCGGATCGAGGGAAGGATTTCAGAAATTC 2700
Db 2641 AACCAACAAACAAATAACAACTCCGGATCGAGGGAAGGATTTCAGAAATTC 2700

AD023594
 ID ADO23594 standard; DNA; 7403 BP.
 XX AC ADO23594;
 XX DT 01-JUL-2004 (first entry)
 XX DE DNA encoding MBP-ToxoP30del2(52-311aa) fusion protein.
 XX KW P30 antigen; Toxo30del13C; Toxo30del2C; ToxoP30 MIX1;
 KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
 KW Toxoplasma gondii; db; gene.
 XX OS Toxoplasma gondii.
 OS Synthetic.
 XX PN US2004067239-A1.
 XX PD 08-APR-2004.
 XX PF 02-OCT-2002; 2002US-00263153.
 XX PR 02-OCT-2002; 2002US-00263153.
 XX PA (MAIN/) MAINE G T.
 PA (PATE/) PATEL C B.
 PA (GINS/) GINSBURG S R.
 PA (BLIE/) BLIESE T R.
 XX PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
 XX DR WPI; 2004-304563/28.
 DR P-PSDB; ADO23595.
 XX PT Novel purified polypeptide having sequence identity to amino acid
 PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
 PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
 PT antibodies to Toxoplasma gondii.
 XX PS Example 2; Fig 11; 114pp; English.
 XX CC The invention relates to a purified P30 antigen (I) chosen from 3 fully
 CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
 CC amino acids added to the C-terminus of the amino acid sequence of
 CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
 CC which at least one of the five C-terminal cysteine amino acids of the
 CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
 CC with alanine, or comprising the amino acid sequence chosen from MBP-
 CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
 CC fusion proteins. (I) is useful for detecting the presence of IgM
 CC antibodies to Toxoplasma gondii in a test sample, which involves
 CC contacting the test sample suspected of containing the IgM antibodies
 CC with a composition comprising (I) and detecting the presence of (I)/IgM
 CC antibody complexes. The present sequence represents DNA encoding a MBP-
 CC ToxoP30 fusion protein of the invention.
 XX SQ Sequence 7403 BP; 1907 A; 1863 C; 1961 G; 1672 T; 0 U; 0 Other;
 Query Match 81.5%; Score 2690.4; DB 12; Length 7403;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CCGACACCATCGAATGTTGCAAAACCTTTCGGGTATGGCATATAGCGCCCGGAAGAGA 60
 DB 1 CCGACACCATCGAATGTTGCAAAACCTTTCGGGTATGGCATATAGCGCCCGGAAGAGA 60
 QY 61 GTCAATTACAGGTGGTGAATGTGAACACAGTATACGATGTGCGAGAGTATGCCG 120
 DB 61 GTCAATTACAGGTGGTGAATGTGAACACAGTATACGATGTGCGAGAGTATGCCG 120
 QY 121 GTGCTCTTATCAGACCGTTTCCGGGTGTGAACACAGCGCCAGCGTTCCTGCCAAA 180
 DB 121 GTGCTCTTATCAGACCGTTTCCGGGTGTGAACACAGCGCCAGCGTTCCTGCCAAA 180

181 CGCGGGAAAAGTGGAAAGCGCGATGGCGAGCTGAATTAATTCCAAACCGCGTGGCAC 240
 DB CGCGGGAAAAGTGGAAAGCGCGATGGCGAGCTGAATTAATTCCAAACCGCGTGGCAC 240
 QY 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGCCCTGC 300
 DB 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGCCCTGC 300
 QY 301 ACGCGCGTGCAGAAATTTGTCCGCGGATTAAATCTCCGCGCGATCAACTGGGTGCCACGC 360
 DB 301 ACGCGCGTGCAGAAATTTGTCCGCGGATTAAATCTCCGCGCGATCAACTGGGTGCCACGC 360
 QY 361 TGGTGGTGTGATGGTAGAAGCGCGCTCGAAGCCTGTAAGCGCGCGGTGCACAATC 420
 DB 361 TGGTGGTGTGATGGTAGAAGCGCGCTCGAAGCCTGTAAGCGCGCGGTGCACAATC 420
 QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACACAGATGCCA 480
 DB 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACACAGATGCCA 480
 QY 481 TTGCTGTGGAAGCTGCCTGCACCTAAATGTTCCGGGTTAATTTCTTGATGTCTCTGACCAGA 540
 DB 481 TTGCTGTGGAAGCTGCCTGCACCTAAATGTTCCGGGTTAATTTCTTGATGTCTCTGACCAGA 540
 QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGTAGCGACTGGGCGTGGAGCATC 600
 DB 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGTAGCGACTGGGCGTGGAGCATC 600
 QY 601 TGGTTCGCAATTTGGTTCACAGCAAAATCGCGTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
 DB 601 TGGTTCGCAATTTGGTTCACAGCAAAATCGCGTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
 QY 661 CGCGTCTGCGCTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
 DB 661 CGCGTCTGCGCTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
 QY 721 CGGAAACGGGAAGCGCACTGGAGTGCCATGTCGGTGTTCACAAACCAACCATGCAATCTCTGA 780
 DB 721 CGGAAACGGGAAGCGCACTGGAGTGCCATGTCGGTGTTCACAAACCAACCATGCAATCTCTGA 780
 QY 781 ATGAGGCAATCGTTCCACCTCGCATCTGTTGCCAACGATCAGATGGCGTGGGGCGCAA 840
 DB 781 ATGAGGCAATCGTTCCACCTCGCATCTGTTGCCAACGATCAGATGGCGTGGGGCGCAA 840
 QY 841 TGGCGCCATTACCGAGTCCGGCTCGCGGTTGGTGGGATATCTCGTATGGGATACG 900
 DB 841 TGGCGCCATTACCGAGTCCGGCTCGCGGTTGGTGGGATATCTCGTATGGGATACG 900
 QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACACCAACCAACAGGATTTTC 960
 DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACACCAACCAACAGGATTTTC 960
 QY 961 GCTGCTGGGGCAACACAGGTGGACCGCTGTGTCGCAACTCTCTCAGGGCCAGGGGTGA 1020
 DB 961 GCTGCTGGGGCAACACAGGTGGACCGCTGTGTCGCAACTCTCTCAGGGCCAGGGGTGA 1020
 QY 1021 AGGGCAATCAGCTGTTGGCCGCTCACTGGTCAAAAGAAAACCAACCAACCAACCAATA 1080
 DB 1021 AGGGCAATCAGCTGTTGGCCGCTCACTGGTCAAAAGAAAACCAACCAACCAACCAATA 1080
 QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCAATTAATGCAAGTGGCAGCAGAGTTT 1140
 DB 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCAATTAATGCAAGTGGCAGCAGAGTTT 1140
 QY 1141 CCGGATGGAACCGGGGAGTGGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
 DB 1141 CCGGATGGAACCGGGGAGTGGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
 QY 1201 GCACAATTTCTATGTTGACAGCTTATCATCGACTGCGGTGCAACCAATGCTTCTGGCG 1260
 DB 1201 GCACAATTTCTATGTTGACAGCTTATCATCGACTGCGGTGCAACCAATGCTTCTGGCG 1260

```
QY 1261 TCAGCGAGCCATCGGAAGCTGTGGTATGGCTGTGCGAGGTGCTAAATCACTGCATAATTTCG 1320
Db 1261 TCAGCGAGCCATCGGAAGCTGTGGTATGGCTGTGCGAGGTGCTAAATCACTGCATAATTTCG 1320
QY 1321 TGTGCTCAAGCGCACCTCCCGTTCCTGATTAATGTTTTTGGCGGACATCATAAACGGTT 1380
Db 1321 TGTGCTCAAGCGCACCTCCCGTTCCTGATTAATGTTTTTGGCGGACATCATAAACGGTT 1380
QY 1381 CTGGCAAAATATTCTGAAATAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGTGGA 1440
Db 1381 CTGGCAAAATATTCTGAAATAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGTGGA 1440
QY 1441 ATTGTGAGCGGATTAACAATTTTACACAGGAACACGCCAGTCCGTTTAGTGTTCACGA 1500
Db 1441 ATTGTGAGCGGATTAACAATTTTACACAGGAACACGCCAGTCCGTTTAGTGTTCACGA 1500
QY 1501 GCACCTTCAACCAAGGACCAATAGATTATGAAACTGAAAGAGGTAAACTCGTAACTCGG 1560
Db 1501 GCACCTTCAACCAAGGACCAATAGCATATGAAATTCGAAAGAGGTAAACTCGTAACTCGG 1560
QY 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATTCGAAAGAGAAATTCGCCACAGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATTCGAAAGAGAAATTCGCCACAGTT 1680
QY 1681 GCGGCAACTGCGGATGCGCCCTGACATTATCTTCTGGGCACACGCCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGCGGATGCGCCCTGACATTATCTTCTGGGCACACGCCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCGGACAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCGGACAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGCTAGCTTACAAACGCAAGCTGATTCCTACCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGCTAGCTTACAAACGCAAGCTGATTCCTACCGATCGCTGTT 1860
QY 1861 GAAGCGTTTATCGCTGATTTATAAAGAGATCTGCTGCCGAACCGCCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTTATCGCTGATTTATAAAGAGATCTGCTGCCGAACCGCCCAAAACCTGGGAA 1920
QY 1921 GAGATCCGCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGCTACTTCACTGGCGGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGCTACTTCACTGGCGGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAARACGCGAAGTACGACATTAAGACGTGGCGGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAARACGCGAAGTACGACATTAAGACGTGGCGGTGGATTAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCACATCGACACGCAAGTGAATTAATGTTGTAACGGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCACATCGACACGCAAGTGAATTAATGTTGTAACGGTACTGCCGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAAACGCCGCCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAAACGCCGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
```

```
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAAGACAAAACCGCTGGGTGCCGTAGCGCTGAAAGTCTTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAAACCGCTGGGTGCCGTAGCGCTGAAAGTCTTTACGAGGAAG 2460
QY 2461 TTGCGGAAAAGATCCACGTTATTTGCCGCCACCATGGAACACGCCCAAGAGGTGAATCATG 2520
Db 2461 TTGCGGAAAAGATCCACGTTATTTGCCGCCACCATGGAACACGCCCAAGAGGTGAATCATG 2520
QY 2521 CCGAAACATCCGCGAGATGTCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAAACATCCGCGAGATGTCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGGTCGTTCAGACTGTCGATGAAGCCCTGAAAGA CGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTCGTTCAGACTGTCGATGAAGCCCTGAAAGA CGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAATAACAATAACAACAACTCGGGATCGAGGGAAGGATTTCAGAATTC 2700
Db 2641 AACAAACAAACAATAACAATAACAACAACTCGGGATCGAGGGAAGGATTTCAGAATTC 2700
```

Search completed: August 1, 2005, 04:28:44
Job time : 1101.6 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 20:22:50 ; Search time 9276 Seconds
(without alignments)
17238.272 Million cell updates/sec

Title: US-09-765-555B-16
Perfect score: 3300
Sequence: 1 ccgacacacgaatggtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hcg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3300	100.0	3300	6	AX202428	Sequence
2	3216.8	97.5	3300	6	AX202430	Sequence
3	3207.2	97.2	3300	6	AX202426	Sequence
4	3207.2	97.2	3300	6	AX202427	Sequence
5	3207.2	97.2	3300	6	AX202429	Sequence
6	2705.8	82.0	7373	6	AX284144	Sequence
7	2705.8	82.0	8101	6	AX172306	Sequence
8	2701	81.8	6648	6	AX377531	Sequence
9	2700	81.8	7475	6	A91965	Sequence 1
10	2700	81.8	7475	6	AR031992	Sequence
11	2700	81.8	7475	6	AR207294	Sequence
12	2698	81.8	9191	6	AX377532	Sequence
13	2676.2	81.1	6724	12	AF097412	Sequence
14	2636.8	79.9	6806	6	AX378208	Sequence
15	2631	79.7	6706	12	AF031088	Sequence
16	2598.2	78.7	6748	12	AF031813	Sequence
17	1609.4	48.8	5558	6	CQ794769	Sequence
18	1474	44.7	4700	6	CQ846980	Sequence
19	1462.4	44.3	4921	6	AX768201	Sequence

20	1462.4	44.3	4935	6	AX768200	Sequence
21	1462.4	44.3	4945	6	AX768202	Sequence
22	1462.4	44.3	4951	6	AX768203	Sequence
c 23	1432.4	43.4	6755	12	AF140577	Integrati
24	1417.6	43.0	1511	6	AX431576	Sequence
25	1417.6	43.0	4297	12	AV219685	Expressio
26	1417.6	43.0	4359	12	AV219686	Expressio
27	1417.6	43.0	4419	12	AV219688	Expressio
28	1417.6	43.0	4855	12	AV219689	Expressio
29	1417.6	43.0	4968	12	AV219687	Expressio
30	1417.6	43.0	6956	12	AV219682	Shuttle e
31	1417.6	43.0	7018	12	AV219683	Shuttle e
32	1417.6	43.0	7509	12	AV219684	Shuttle e
33	1400.8	42.4	4566	12	AF177933	Cloning v
34	1400.8	42.4	4635	12	AF177932	Cloning v
35	1400.8	42.4	6399	12	AY302759	Expressio
36	1400.8	42.4	6587	12	AY289542	Expressio
37	1362	41.3	5201	6	AR493834	Sequence
38	1362	41.3	5201	6	AR493835	Sequence
39	1329.2	40.3	4806	12	AF050464	Expressio
c 40	1238.4	37.5	4803	12	AY243506	Cloning v
c 41	1238.4	37.5	5024	6	CQ846977	Sequence
42	1237.2	37.5	1922	6	AX478528	Sequence
43	1204	36.5	4506	12	AB116367	Expressio
44	1204	36.5	4557	6	EL3740	DNA sequenc
45	1204	36.5	4557	6	AR373251	Sequence

ALIGNMENTS

RESULT 1
AX202428
LOCUS AX202428 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 16 from Patent WO0152620.
ACCESSION AX202428
VERSION AX202428.1 GI:15392176
KEYWORDS synthetic construct
SOURCE other sequences; artificial sequences.
ORGANISM
REFERENCE 1
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 16 26-JUL-2001;
The Scripps Research Institute (US) ; SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES
source Location/Qualifiers
1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-m3 and zinc finger protein ZFPm3"

ORIGIN
Query Match 100.0%; Score 3300; DB 6; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CGACACCATGATGGTGGCAAAACCTTTCGGGTATGCGCATGATAGCCCGGAAGAGA 60
DB 1 CGACACCATGATGGTGGCAAAACCTTTCGGGTATGCGCATGATAGCCCGGAAGAGA 60
QY 61 GTCATTCAGGTGGTGAATGTGAACCAAGTAACTTATACGATGTCGACAGTATGCCG 120
DB 61 GTCATTCAGGTGGTGAATGTGAACCAAGTAACTTATACGATGTCGACAGTATGCCG 120
QY 121 GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGTAACTTATACGATGTCGACAGTATGCCG 180
DB 121 GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGTAACTTATACGATGTCGACAGTATGCCG 180
QY 181 CGCGGAAAAAGTGGAAACGCGGATGGCGGAGCTGAATTACATCCCAACCGGTGGCAC 240

Db	191	CGCGGAAAGTGGAGCGGCGATGGCGAGGCTGAAATTACAAATCCCAACCGCTGGCAC	240
Qy	241	AACAACTGGCGGCAAAACAGTCTGTTCTGATTTGGGCTTGGCCACCTCCAGTCTGGGCCCTGC	300
Db	241	AACAACTGGCGGCAAAACAGTCTGTTCTGATTTGGCGTTGGCCACCTCCAGTCTGGGCCCTGC	300
Qy	301	ACGGCCGTCGCAAAATTTGTCGGCGCGATTAATCTCGCGCCGATCAACTGCGGTGCGACGG	360
Db	301	ACGGCCGTCGCAAAATTTGTCGGCGCGATTAATCTCGCGCCGATCAACTGCGGTGCGACGG	360
Qy	361	TGGTGGTTCGATGGTAGAAAGAAAGCGGCTCGAAGCCTGTAAAGCGGCGGTGCACAATC	420
Db	361	TGGTGGTTCGATGGTAGAAAGAAAGCGGCTCGAAGCCTGTAAAGCGGCGGTGCACAATC	420
Qy	421	TTCTCGCGCAACGCGTCAGTGGCGTGATCAATTAATATCCCGCTGGATGACCAAGGATGCCA	480
Db	421	TTCTCGCGCAACGCGTCAGTGGCGTGATCAATTAATATCCCGCTGGATGACCAAGGATGCCA	480
Qy	481	TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGTGACCGA	540
Db	481	TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGTGACCGA	540
Qy	541	CACCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACGCGACTGGCGCGTGGAGCATC	600
Db	541	CACCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACGCGACTGGCGCGTGGAGCATC	600
Qy	601	TGCTCGCAATGGGTCACAGCAAAATCGCGCTGTAGCGGGGCCCATTAAGTTCTGCTCGG	660
Db	601	TGCTCGCAATGGGTCACAGCAAAATCGCGCTGTAGCGGGGCCCATTAAGTTCTGCTCGG	660
Qy	661	CGCTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGCGATAG	720
Db	661	CGCTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGCGATAG	720
Qy	721	CGGAACGGGAAGCGACTGGAGTGCCATGTCGGCTTTTCAACAAACCATGCAAAATGCTGA	780
Db	721	CGGAACGGGAAGCGACTGGAGTGCCATGTCGGCTTTTCAACAAACCATGCAAAATGCTGA	780
Qy	781	ATGAGGCAATCGTTTCCCACTGCGATGTGTTGTCACACGATCAGATGGCGCTGGCGCGAA	840
Db	781	ATGAGGCAATCGTTTCCCACTGCGATGTGTTGTCACACGATCAGATGGCGCTGGCGCGAA	840
Qy	841	TGCGCGCATTTACCGAGTCCGGGCTCGCGTTGGTGGCGGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGCGCATTTACCGAGTCCGGGCTCGCGTTGGTGGCGGATATCTCGGTAGTGGGATACG	900
Qy	901	ACGATACCGAAGACAGCTCATGTTTATATCCGCGCTTAAACACCATCAAAACAGATTTTC	960
Db	901	ACGATACCGAAGACAGCTCATGTTTATATCCGCGCTTAAACACCATCAAAACAGATTTTC	960
Qy	961	GCCTGTGGGGCAAAACAGGCTGGAACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020
Db	961	GCCTGTGGGGCAAAACAGGCTGGAACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020
Qy	1021	AGGGCAATCAGCTGTTCGCCGTCTCACTGGTGAAGAAAGAAACACACCTCGCGGCCAATA	1080
Db	1021	AGGGCAATCAGCTGTTCGCCGTCTCACTGGTGAAGAAAGAAACACACCTCGCGGCCAATA	1080
Qy	1081	CGAAACCGGCTCTCCCGCGGCTTGGCCGATTCATTAATGCAAGCTGGCACGACAGGTTT	1140
Db	1081	CGAAACCGGCTCTCCCGCGGCTTGGCCGATTCATTAATGCAAGCTGGCACGACAGGTTT	1140
Qy	1141	CCGCACTGGAAGCGGCGAGTGAAGCAACGCAATTAATGTGAGTAGCTCACTCATTTAG	1200
Db	1141	CCGCACTGGAAGCGGCGAGTGAAGCAACGCAATTAATGTGAGTAGCTCACTCATTTAG	1200
Qy	1201	GCAAAATTCATGTTTGAAGCTTATCATCGACTGCAAGCTGCAAAATGCTTCTCGCG	1260
Db	1201	GCAAAATTCATGTTTGAAGCTTATCATCGACTGCAAGCTGCAAAATGCTTCTCGCG	1260
Qy	1261	TCAGGCAAGCATCGAAGCTGGTATGGCTGTGAGGTGTTAAATCACTGCATAAATTCG	1320

Db	1261	TCAGCAGCCATCGAAAGCTGTGGTATGGCTGTGCAGGTCGTAATCACTGCATAATTCG	1320
Qy	1321	TGTCGCTCAAGCGCACCTCCGGTCTCGGATAATGTTTTTTCGCGCGACATCAACCGGTT	1380
Db	1321	TGTGCTCAAGCGCACCTCCGGTCTCGGATAATGTTTTTTCGCGCGACATCAACCGGTT	1380
Qy	1381	CTGCAAAATATCTGAAATGAGCTGTTGTGACAAATTAATCATCGGCTCGTAAATGTGTGGA	1440
Db	1381	CTGCAAAATATCTGAAATGAGCTGTTGTGACAAATTAATCATCGGCTCGTAAATGTGTGGA	1440
Qy	1441	ATTGTGAGCGGATAACAATTTTCACACAGGAAAACAGCCAGTCGGTTAGTGTGTTTTCA	1500
Db	1441	ATTGTGAGCGGATAACAATTTTCACACAGGAAAACAGCCAGTCGGTTAGTGTGTTTTCA	1500
Qy	1501	GCATTTCAACAAACAAAGGACCATAGATTATGAAAACCTGAAGAAGGTAACCTGGTAAATCTCG	1560
Db	1501	GCATTTCAACAAACAAAGGACCATAGATTATGAAAACCTGAAGAAGGTAACCTGGTAAATCTCG	1560
Qy	1561	ATTAAACGGGATAAAGGCTATAACCGTCTCGCTGAAAGTCGGTAAGAAAATTCGAGAAAGAT	1620
Db	1561	ATTAAACGGGATAAAGGCTATAACCGTCTCGCTGAAAGTCGGTAAGAAAATTCGAGAAAGAT	1620
Qy	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAACTCGAAGAGAAAATCCCAACAGGTT	1680
Db	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAACTCGAAGAGAAAATCCCAACAGGTT	1680
Qy	1681	CGGCAAACTCGGAGATGGCCCTGACATATCTCTCGGCAACGACCGCTTTGGTGCTAC	1740
Db	1681	CGGCAAACTCGGAGATGGCCCTGACATATCTCTCGGCAACGACCGCTTTGGTGCTAC	1740
Qy	1741	GCTCAATCTGGCCTGTGGCTGAAATCACCCGGACAAAGCGTCCAGACACAGCTGAT	1800
Db	1741	GCTCAATCTGGCCTGTGGCTGAAATCACCCGGACAAAGCGTCCAGACACAGCTGAT	1800
Qy	1801	CGGTTTACCTGGGATGCGGTACGTTTACAACGGCAAGCTGATTTGCTTACCAGTACGCTGT	1860
Db	1801	CGGTTTACCTGGGATGCGGTACGTTTACAACGGCAAGCTGATTTGCTTACCAGTACGCTGT	1860
Qy	1861	GAAGCGTTATCGCTGATTTATACAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA	1920
Db	1861	GAAGCGTTATCGCTGATTTATACAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA	1920
Qy	1921	GAGATCCGCGCTGGATAAAGAACTGAAGAGCGGAAAGGTAAGAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCGCGCTGGATAAAGAACTGAAGAGCGGAAAGGTAAGAGCGCGCTGATGTTCAAC	1980
Qy	1981	CTGCAAGAACCGTACTTCACTGGCCGCTGATTTGCTGTGACGGGGGTTATGCGTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTCACTGGCCGCTGATTTGCTGTGACGGGGGTTATGCGTTCAAG	2040
Qy	2041	TATGAAAAACGGCAAGTACGACATTAAGAAGCTGGGCGTGGATAACGCTGGCGCGAAAGCG	2100
Db	2041	TATGAAAAACGGCAAGTACGACATTAAGAAGCTGGGCGTGGATAACGCTGGCGCGAAAGCG	2100
Qy	2101	GGTCTGACCTTCTGTTGTGACCTGATTAATAACAAACACATGAATGCGAGACACCGATTAC	2160
Db	2101	GGTCTGACCTTCTGTTGTGACCTGATTAATAACAAACACATGAATGCGAGACACCGATTAC	2160
Qy	2161	TCCATCGCAGAGCTGCCCTTTTAATAAGGCGAAAACAGCGATGACCATCAACGGGCCCGTGG	2220
Db	2161	TCCATCGCAGAGCTGCCCTTTTAATAAGGCGAAAACAGCGATGACCATCAACGGGCCCGTGG	2220
Qy	2221	GCATGTCCAAACATCGACACCGGATAAGTATGTTGTTAAACGGTACTGCCGACCTTC	2280
Db	2221	GCATGTCCAAACATCGACACCGGATAAGTATGTTGTTAAACGGTACTGCCGACCTTC	2280
Qy	2281	AAGGGTCAACCATCCAAACCGTTTCTGTGCGTGTGAGCGCAGGTATTAAACGCGCCAGT	2340
Db	2281	AAGGGTCAACCATCCAAACCGTTTCTGTGCGTGTGAGCGCAGGTATTAAACGCGCCAGT	2340
Qy	2341	CGAAACAAAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGCTCTG	2400
Db	2341	CGAAACAAAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGCTCTG	2400

Qy	2401	GAAGCGGTTAATAAGACAAACCCGTCGGTGCCTAGCGCTGAAGTCTTACGAGGAAG	2460
Db	2401	GAAGCGGTTAATAAGACAAACCCGTCGGTGCCTAGCGCTGAAGTCTTACGAGGAAG	2460
Qy	2461	TTGGCGAAGATCCACGCTATTGCGCCACCACTGGAAAGCGCCACAGAAAGTGAATCATG	2520
Db	2461	TTGGCGAAGATCCACGCTATTGCGCCACCACTGGAAAGCGCCACAGAAAGTGAATCATG	2520
Qy	2521	CCGAACATCCCGCAGATGTCGCGCTTTCTGGTATGCGCTGCTACTGCGGTGATCAACGCC	2580
Db	2521	CCGAACATCCCGCAGATGTCGCGCTTTCTGGTATGCGCTGCTACTGCGGTGATCAACGCC	2580
Qy	2581	GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG	2640
Db	2581	GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG	2640
Qy	2641	AACACACACAAATAACAATAACAAACCTCGGGATCGAGGAGGATTCGAAATTC	2700
Db	2641	AACACACACAAATAACAATAACAAACCTCGGGATCGAGGAGGATTCGAAATTC	2700
Qy	2701	GGATCCTCTCTCTGTCGGCCAGCGGCCCTCGAGCCCGGGGAGAACCTTATCTTGT	2760
Db	2701	GGATCCTCTCTCTGTCGGCCAGCGGCCCTCGAGCCCGGGGAGAACCTTATCTTGT	2760
Qy	2761	CCGGAATGTGTAAGTCTCTCAGCGATCCTGSCCACTTGTTGCGCCACGAGCGTACCCAC	2820
Db	2761	CCGGAATGTGTAAGTCTCTCAGCGATCCTGSCCACTTGTTGCGCCACGAGCGTACCCAC	2820
Qy	2821	ACGGGTGAAAAACCTATTAATGCCAGATGCGGCAAACTTTTAGCACAGCGGCTCC	2880
Db	2821	ACGGGTGAAAAACCTATTAATGCCAGATGCGGCAAACTTTTAGCACAGCGGCTCC	2880
Qy	2881	CTGGTGCCTCAACAGCACTCATACTGCGGAGAGCCATACAAATGTCAGAAATGTGCG	2940
Db	2881	CTGGTGCCTCAACAGCACTCATACTGCGGAGAGCCATACAAATGTCAGAAATGTGCG	2940
Qy	2941	AAGTCTTTTCAGCAGAGTCCAGCTGTGTGCGCCACCAACGTACTCACACCGGGGAGAG	3000
Db	2941	AAGTCTTTTCAGCAGAGTCCAGCTGTGTGCGCCACCAACGTACTCACACCGGGGAGAG	3000
Qy	3001	CCCTATGCTTGTCCCGAATGTTGAAGTCCCTTACGCCAGAGAGCTCCCTGTCGCCAC	3060
Db	3001	CCCTATGCTTGTCCCGAATGTTGAAGTCCCTTACGCCAGAGAGCTCCCTGTCGCCAC	3060
Qy	3061	CAGCGTACCCACACCGGTGAAAAACCGTATAATGCCAGAGTGCAGCAAAATCTTTTGT	3120
Db	3061	CAGCGTACCCACACCGGTGAAAAACCGTATAATGCCAGAGTGCAGCAAAATCTTTTGT	3120
Qy	3121	GACTGCCGCGACTTTGCTCGCCATCAACGCACTCATACTGCGGAGAGCCATACAAATGT	3180
Db	3121	GACTGCCGCGACTTTGCTCGCCATCAACGCACTCATACTGCGGAGAGCCATACAAATGT	3180
Qy	3181	CCAGAAATGGAAGTCTTTCTCCAAATCCAGCCATCTGTCGGGACCAACAGTACTCAC	3240
Db	3181	CCAGAAATGGAAGTCTTTCTCCAAATCCAGCCATCTGTCGGGACCAACAGTACTCAC	3240
Qy	3241	ACCGGTAAAAAATAGTGGCCAGCGCCAGTACCGTACGAGTTCGCGACTACGCT	3300
Db	3241	ACCGGTAAAAAATAGTGGCCAGCGCCAGTACCGTACGAGTTCGCGACTACGCT	3300

RESULT 2

AX202430

LOCUS

AX202430

Sequence 18 from Patent W00152620.

DEFINITION

AX202430

ACCESSION

AX202430.1

GI:15392178

VERSION

KEYWORDS

SOURCE

synthetic construct

ORGANISM

synthetic construct

other sequences; artificial sequences.

REFERENCE

1

AUTHORS	Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.									
TITLE	Methods and compositions to modulate expression in plants									
JOURNAL	Patent: WO 0152620-A 18 26-JUL-2001; The Scripps Research Institute (US) ; SYNGENTA AGRICULTURAL DISCOVERY, INC. (CA)									
FEATURES	Location/Qualifiers									
source	1..3300 /organism="synthetic construct" /mol_type="unassigned DNA" /db_xref="taxon:32630" /notes="Partial sequence of pMal-Ap3 and zinc finger protein ZFPp3"									
ORIGIN										
Query Match	97.5%; Score 3216.8; DB 6; Length 3300;									
Best Local Similarity	98.4%; Pred. No. 0;									
Matches 3248; Conservative	0; Mismatches 52; Indels 0; Gaps 0;									
Qy	1	CCGACACCATCGAATGGTGC	AAAACCTTTTCGCGGTATGGCATGATAGCCCGCCGGA	GAGA	60					
Db	1	CCGACACCATCGAATGGTGC	AAAACCTTTTCGCGGTATGGCATGATAGCCCGCCGGA	GAGA	60					
Qy	61	GTCAATTTCAGGGTGGTGA	TGTGAAACCAAGTAACTGATACGATGTCGCAGAGTATGCCG	120						
Db	61	GTCAATTTCAGGGTGGTGA	TGTGAAACCAAGTAACTGATACGATGTCGCAGAGTATGCCG	120						
Qy	121	GTGTCCTTTATCAGACCCG	TTTCCCGCGTGTGTAAACACGAGCCAGCCACCGTTTCTTCGCGAAA	180						
Db	121	GTGTCCTTTATCAGACCG	TTTCCCGCGTGTGTAAACACGAGCCAGCCACCGTTTCTTCGCGAAA	180						
Qy	181	CGCGGGAAGAAAGTGAAG	CGCGGCGATGGCGAGCTGAAATTACATTTCCCAACCGCGGTGGCAC	240						
Db	181	CGCGGGAAGAAAGTGAAG	CGCGGCGATGGCGAGCTGAAATTACATTTCCCAACCGCGGTGGCAC	240						
Qy	241	AACAACCTGGCGGCAACAG	CTGCTGTAATGGCGTTGGCACCTCCAGTCTGGCCCTGCG	300						
Db	241	AACAACCTGGCGGCAACAG	CTGCTGTAATGGCGTTGGCACCTCCAGTCTGGCCCTGCG	300						
Qy	301	ACGCGCGCTCGCAAAAT	TGTCCGCGCGATTAAATCTCCGCGCGATCAACTGGGTGCCAGCG	360						
Db	301	ACGCGCGCTCGCAAAAT	TGTCCGCGCGATTAAATCTCCGCGCGATCAACTGGGTGCCAGCG	360						
Qy	361	TGTTGTTGTCGATGTGTA	AGCGAGCGGCTCGAAGCTGTGTAAGCGGCGGTGCACAATC	420						
Db	361	TGTTGTTGTCGATGTGTA	AGCGAGCGGCTCGAAGCTGTGTAAGCGGCGGTGCACAATC	420						
Qy	421	TTCTCGCGCAACGCGTCA	GTTGGGCTGATCAATTAACCTATCCGCTGGATGACACAGGATGCCA	480						
Db	421	TTCTCGCGCAACGCGTCA	GTTGGGCTGATCAATTAACCTATCCGCTGGATGACACAGGATGCCA	480						
Qy	481	TTGCTGTGGAAGCTGCTG	CACCTAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCCAGA	540						
Db	481	TTGCTGTGGAAGCTGCTG	CACCTAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCCAGA	540						
Qy	541	CACCCATCAACAGTATTAT	TTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600						
Db	541	CACCCATCAACAGTATTAT	TTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600						
Qy	601	TGTTGCGGATTTGGGTCA	CACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCGTCTCGG	660						
Db	601	TGTTGCGGATTTGGGTCA	CACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCGTCTCGG	660						
Qy	661	CGCGTCTGCGTCTGGCTG	GCGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
Db	661	CGCGTCTGCGTCTGGCTG	GCGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
Qy	721	CGGAACGGGAAGGCGACT	GAGTGGCCATGTCGGTTCACAAACCAATGCAAAATGCTGA	780						
Db	721	CGGAACGGGAAGGCGACT	GAGTGGCCATGTCGGTTCACAAACCAATGCAAAATGCTGA	780						
Qy	781	ATGAGGCGATCGTTCCCA	CTGCGATGTTGCAACAGATCAGATGGCGCTGGGCGCAA	840						
Db	781	ATGAGGCGATCGTTCCCA	CTGCGATGTTGCAACAGATCAGATGGCGCTGGGCGCAA	840						

QY 841 TCGCGCCATTACCGAGTCCGGCTCGCGTTGGTGGGATATCTCGTAGTGGGATACG 900
Db |||||
QY 841 TCGCGCCATTACCGAGTCCGGCTCGCGTTGGTGGGATATCTCGTAGTGGGATACG 900
Db |||||
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCGCCGTTAAACCACTCAAAACAGGATTTTC 960
Db |||||
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCGCCGTTAAACCACTCAAAACAGGATTTTC 960
Db |||||
QY 961 GCCTCTGGGGCAAAACACGAGTGGACCGCTTGCTGCAACTCTCTCAGGCGCAGCGGTGA 1020
Db |||||
QY 961 GCCTCTGGGGCAAAACACGAGTGGACCGCTTGCTGCAACTCTCTCAGGCGCAGCGGTGA 1020
Db |||||
QY 1021 AGGCAATCAGCTGTGTGGCCGCTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
Db |||||
QY 1021 AGGCAATCAGCTGTGTGGCCGCTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
Db |||||
QY 1081 CGCAACCCCTCTCCCGCGGTTGGCGGATTCATTAATGAGTGGCGACGACGAGTTT 1140
Db |||||
QY 1081 CGCAACCCCTCTCCCGCGGTTGGCGGATTCATTAATGAGTGGCGACGACGAGTTT 1140
Db |||||
QY 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGAGTGGCGACGACGATAG 1200
Db |||||
QY 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGAGTGGCGACGACGATAG 1200
Db |||||
QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAACGCTGACCAATGCTTCTGGCG 1260
Db |||||
QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAACGCTGACCAATGCTTCTGGCG 1260
Db |||||
QY 1261 TCAGGACGCAATCGGAAGCTGTGATGCTGTGAGAGTGAACCAATGCTGCAATTCG 1320
Db |||||
QY 1261 TCAGGACGCAATCGGAAGCTGTGATGCTGTGAGAGTGAACCAATGCTGCAATTCG 1320
Db |||||
QY 1321 TGTGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTTCGCGCGACATCAACCGTT 1380
Db |||||
QY 1321 TGTGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTTCGCGCGACATCAACCGTT 1380
Db |||||
QY 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTTGGA 1440
Db |||||
QY 1441 ATTGTGAGCGGATTAACAAATTTCAACAGGAACAGCCAGTCCGTTTAGTGTGTTTCAGGA 1500
Db |||||
QY 1441 ATTGTGAGCGGATTAACAAATTTCAACAGGAACAGCCAGTCCGTTTAGTGTGTTTCAGGA 1500
Db |||||
QY 1501 GCACCTTCAACCAAGGACCATAGATTAATGAAACTGAAGAGTAAACTGGTAATCTGG 1560
Db |||||
QY 1501 GCACCTTCAACCAAGGACCATAGATTAATGAAACTGAAGAGTAAACTGGTAATCTGG 1560
Db |||||
QY 1561 ATTAACGGCGATTAAGGCTATAACGGTCTCGCTGAAFTCGGTAAGAAATTCGAGAAAGAT 1620
Db |||||
QY 1561 ATTAACGGCGATTAAGGCTATAACGGTCTCGCTGAAFTCGGTAAGAAATTCGAGAAAGAT 1620
Db |||||
QY 1621 ACCGGAATTAAGTCAACCGTTGACATCCGGATTAACGGAGAGAAATTCACAGGTT 1680
Db |||||
QY 1621 ACCGGAATTAAGTCAACCGTTGACATCCGGATTAACGGAGAGAAATTCACAGGTT 1680
Db |||||
QY 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTGGGCAACAGCAGCTTGGTGGCTAC 1740
Db |||||
QY 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTGGGCAACAGCAGCTTGGTGGCTAC 1740
Db |||||
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCCCGGACAAAGGTTTCAGGACAAAGCTGTAT 1800
Db |||||
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCCCGGACAAAGGTTTCAGGACAAAGCTGTAT 1800
Db |||||
QY 1801 CCGTTTACCTGGGATGCGTAGCTTCAACGGCAAGCTGATTCGTTACCCGATCGCTGTT 1860
Db |||||
QY 1801 CCGTTTACCTGGGATGCGTAGCTTCAACGGCAAGCTGATTCGTTACCCGATCGCTGTT 1860
Db |||||
QY 1861 GAAGCGTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAACCCGCAAAACCTGGGAA 1920
Db |||||
QY 1861 GAAGCGTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAACCCGCAAAACCTGGGAA 1920
Db |||||

QY 1921 GAGATCCGGCGCTGGATTAAGAACTGAAGCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
Db |||||
QY 1921 GAGATCCGGCGCTGGATTAAGAACTGAAGCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
Db |||||
QY 1981 CTGCAAGAAACCGTACTTTCACCTGCGCGCTGANTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db |||||
QY 1981 CTGCAAGAAACCGTACTTTCACCTGCGCGCTGANTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db |||||
QY 2041 TATGAAAAACCGCAAGTAGTACGACATTAAGAGCGTGGCGTGGATACGCTGGCGCAAGAGCG 2100
Db |||||
QY 2041 TATGAAAAACCGCAAGTAGTACGACATTAAGAGCGTGGCGTGGATACGCTGGCGCAAGAGCG 2100
Db |||||
QY 2101 GGTCTGACCTTCTCGTGGTTGACCTGATTAAGAAACAAACACATGATGACAGACCGATTC 2160
Db |||||
QY 2101 GGTCTGACCTTCTCGTGGTTGACCTGATTAAGAAACAAACACATGATGACAGACCGATTC 2160
Db |||||
QY 2161 TCCATCGCAAGAGTGCCTTTTAAAGAGCGAAACAGCGATGACCATCAACGGCGCGGTGG 2220
Db |||||
QY 2161 TCCATCGCAAGAGTGCCTTTTAAAGAGCGAAACAGCGATGACCATCAACGGCGCGGTGG 2220
Db |||||
QY 2221 GCATGGTCCAAACATCGACACAGCAAGTGAATTAATGCTGTAAACGCTGACGACCTTC 2280
Db |||||
QY 2221 GCATGGTCCAAACATCGACACAGCAAGTGAATTAATGCTGTAAACGCTGACGACCTTC 2280
Db |||||
QY 2281 AAGGTCACACATCCAAACCGTTTGGCGTGTGAGCGCAGGTATTAACGGCGCGAGT 2340
Db |||||
QY 2281 AAGGTCACACATCCAAACCGTTTGGCGTGTGAGCGCAGGTATTAACGGCGCGAGT 2340
Db |||||
QY 2341 CCGAAACAAAGAGTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
Db |||||
QY 2341 CCGAAACAAAGAGTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
Db |||||
QY 2401 GAAGCGTTTAAAGACAAACCGTGGTGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Db |||||
QY 2401 GAAGCGTTTAAAGACAAACCGTGGTGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Db |||||
QY 2461 TTGGCAAGAGTCCAGTATTTGCGCCACCATCGAAACCGCCAGAAAGTGAATCATG 2520
Db |||||
QY 2461 TTGGCAAGAGTCCAGTATTTGCGCCACCATCGAAACCGCCAGAAAGTGAATCATG 2520
Db |||||
QY 2521 CCGAAACATCCGAGATGTCGGCTTTCTGGTATGCGCTGGTACTGCGGTGATCAAGCC 2580
Db |||||
QY 2521 CCGAAACATCCGAGATGTCGGCTTTCTGGTATGCGCTGGTACTGCGGTGATCAAGCC 2580
Db |||||
QY 2581 GCCAGCGCTGTCAGACTGTCGATGAAGCCCTGAAAGACCGCGCAGACTAATTCGAGCTG 2640
Db |||||
QY 2581 GCCAGCGCTGTCAGACTGTCGATGAAGCCCTGAAAGACCGCGCAGACTAATTCGAGCTG 2640
Db |||||
QY 2641 AACCAACAAACAAATTAACAAATTAACAACTCGGATCGAGGAAAGATTTTCAGAAATTC 2700
Db |||||
QY 2641 AACCAACAAACAAATTAACAAATTAACAACTCGGATCGAGGAAAGATTTTCAGAAATTC 2700
Db |||||
QY 2701 GGATCTCTTCTCTGTCGCGCAGCGGCTCTGAGCCCGGGGAGAGCCCTATGCTGTG 2760
Db |||||
QY 2701 GGATCTCTTCTCTGTCGCGCAGCGGCTCTGAGCCCGGGGAGAGCCCTATGCTGTG 2760
Db |||||
QY 2761 CCGGAATGTTGAAGTCTTTCAGCGATCTTGGCCACTGCTGGTCCGCAACGAGCTACCCAC 2820
Db |||||
QY 2761 CCGGAATGTTGAAGTCTTTCAGCGATCTTGGCCACTGCTGGTCCGCAACGAGCTACCCAC 2820
Db |||||
QY 2821 ACGGGTGAAGAAACCGTATTAATGCGCAGAGTGGGCAAAATCTTTTAGCACACCGGCTCC 2880
Db |||||
QY 2821 ACGGGTGAAGAAACCGTATTAATGCGCAGAGTGGGCAAAATCTTTTAGCACACCGGCTCC 2880
Db |||||
QY 2881 CTGGTGGCGATCAACGCACTCATCTGCGGAGAGCCCATCAAAATGTCAGAAATGTGGC 2940
Db |||||
QY 2881 CTGGTGGCGATCAACGCACTCATCTGCGGAGAGCCCATCAAAATGTCAGAAATGTGGC 2940
Db |||||
QY 2941 AAGTCTTTTCAGCGAGCTCCAGCTGTCGCGCCACCAACGCTACTCACACCGGGGAGAG 3000
Db |||||
QY 2941 AAGTCTTTTCAGCGAGCTCCAGCTGTCGCGCCACCAACGCTACTCACACCGGGGAGAG 3000
Db |||||
QY 3001 CCCTATGCTGTGTCGGAAATGTGTTAAGTCTTTTCAGCCAGAGCAGCTCCCTGTGTCGCCAC 3060
Db |||||

Db 3001 |||||||CCCTATGCTTGTCCGGAATGTTGAAGTCTCTTTCAGCACACGAGTGGCTCTCTTGGTTAGACAC 3060
Qy 3061 CAGCGTACCCACACACGGGTGAAAAACCGGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGT 3120
Db 3061 CAGCGTACCCACACACGGGTGAAAAACCGGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
Qy 3121 GACTGCCGGACCTTGTCTGGCCATCAACGCACTCATCTACTGCGGAGAACCATACAAATGT 3180
Db 3121 CAGCGCGCCACCTTGAAACGCCCATCAAGGCATCTACTGCGGAGAACCATACAAATGT 3180
Qy 3181 CCAGAATGTGCAAGTCTTCTCCCAATCCAGCCATCTCGTCCGCGACCAAGTACTCAC 3240
Db 3181 CCAGAATGTGCAAGTCTTCTCAACTTCAGGCAACTTGGTCCGTCACCAAGTACTCAC 3240
Qy 3241 ACCCGTAAAAAACTAGTGGCCAGCGCCGACAGTACCGGTAGCAGTTCGCGACTACGCT 3300
Db 3241 ACCCGTAAAAAACTAGTGGCCAGCGCCGACGACGCTTCCGACTACGCT 3300

RESULT 3
AX202426
LOCUS AX202426 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 14 from Patent WO0152620.
ACCESSION AX202426
VERSION AX202426.1 GI:15392173
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 14 26-JUL-2001;
The Scripps Research Institute (US) ; SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)

FEATURES
source Location/Qualifiers
1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-m1 and zinc finger protein
ZFpml"

ORIGIN
Query Match 97.2%; Score 3207.2; DB 6; Length 3300;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
Qy 61 GTCATTCAGGGTGTGAATGTGAACACCAAGTATACGATTCGAGAGTATGCG 120
Db 61 GTCATTCAGGGTGTGAATGTGAACACCAAGTATACGATTCGAGAGTATGCG 120
Qy 121 GTGTCCTTATCAGACCGGTTTCCCGCGTGTGAACAGCGACGACGCTTCTCGGAAAA 180
Db 121 GTGTCCTTATCAGACCGGTTTCCCGCGTGTGAACAGCGACGACGCTTCTCGGAAAA 180
Qy 181 CCGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAAATTACATTCCTCAACCGCGTGGC 240
Db 181 CCGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAAATTACATTCCTCAACCGCGTGGC 240
Qy 241 AACAACTGGCGGGCAAAACAGTCTGTGCTGATGGCGGTGGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGCAAAACAGTCTGTGCTGATGGCGGTGGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGCCCGTTCGCAAAATGTTCGCGGGGATTAATCTCGCGCCGATCAACTGGGTGCCAGG 360
Db 301 ACGCCCGTTCGCAAAATGTTCGCGGGGATTAATCTCGCGCCGATCAACTGGGTGCCAGG 360

Qy 361 TGGTGGTGTCTGATGGTAGAACGAAGCGCGTGAAGCCTGTAAAGCGCGGTGCACAATC 420
Db 361 TGGTGGTGTCTGATGGTAGAACGAAGCGCGTGAAGCCTGTAAAGCGCGGTGCACAATC 420
Qy 421 TTCTCGGCAACGGCTCAGTGGGCTGATCAATACTATCCGTGATCCAGTACCAAGATGCCA 480
Db 421 TTCTCGGCAACGGCTCAGTGGGCTGATCAATACTATCCGTGATCCAGTACCAAGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTCTGACCA 540
Db 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTCTGACCA 540
Qy 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTGCATTCGGGTACACAGCAAAATCGGCGTGTAGCGGGCCCATTAAGTGTCTGCTCG 660
Db 601 TGGTGCATTCGGGTACACAGCAAAATCGGCGTGTAGCGGGCCCATTAAGTGTCTGCTCG 660
Qy 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGCGCACTGGAGTGCATGTCGGTTCCTCAACAAACCATGCAATGCTGA 780
Db 721 CGGAACGGGAAGCGCACTGGAGTGCATGTCGGTTCCTCAACAAACCATGCAATGCTGA 780
Qy 781 ATGAGGGCATGTTCCCACTCGCATGCTGTTGTCACACGATCAGATGCGCTGGCGCGAA 840
Db 781 ATGAGGGCATGTTCCCACTCGCATGCTGTTGTCACACGATCAGATGCGCTGGCGCGAA 840
Qy 841 TGGCGCCATTACCGAGTCCGGGCTGCGGTGTGGTGCAGATATCTCGTAGTGGGATAG 900
Db 841 TGGCGCCATTACCGAGTCCGGGCTGCGGTGTGGTGCAGATATCTCGTAGTGGGATAG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACCAACCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACCAACCATCAACAGGATTTTC 960
Qy 961 GCCTGTGGGGCAACACAGCGTGGACCGCTTGTGTCGCAACTCTCTCAGGGCGCGCGTGA 1020
Db 961 GCCTGTGGGGCAACACAGCGTGGACCGCTTGTGTCGCAACTCTCTCAGGGCGCGCGTGA 1020
Qy 1021 AGGCAATCAGCTGTGTCCTCTCACTGTGTAAGAAAAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGTCCTCTCACTGTGTAAGAAAAAACCAACCTGGCGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGGACGACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGGACGACAGGTTT 1140
Qy 1141 CCGGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTGTAGTCTACTCATTTAG 1200
Db 1141 CCGGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTGTAGTCTACTCATTTAG 1200
Qy 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGTGCACCAATGCTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGTGCACCAATGCTCTGGCG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGGTATGCGTGTGAGCTGTAAATCATCTGATTAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGCGTGTGAGCTGTAAATCATCTGATTAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACCTCCCGTCTCGGATAATGTTTTTGGCGCGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACCTCCCGTCTCGGATAATGTTTTTGGCGCGACATCATACGGTT 1380
Qy 1381 CTGGCAAAATATCTCGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTTGA 1440
Db 1381 CTGGCAAAATATCTCGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTTGA 1440
Qy 1441 ATTGTAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500

Db 1441 ATTGTGAGCGGATAACAATTTTACACAGAGAAACAGCCAGTCGGTTTAGTGTTTTCAGCA 1500
Qy 1501 GCACITTCACCAACAGGACCAATAGATTATGAAAATCTGAAGAGGTAAACTGTTATCTGG 1560
Db 1501 GCACITTCACCAACAGGACCAATAGATTATGAAAATCTGAAGAGGTAAACTGTTATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACCGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACCGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCAGGTT 1680
Qy 1681 GCGGCAACTGCGGATGCGGCTTGACATTAATCTCTGGGCACACGCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGCGGATGCGGCTTGACATTAATCTCTGGGCACACGCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTAGTTTAAACCGGCAAGCTGATGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTAGTTTAAACCGGCAAGCTGATGCTTACCCGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCCGACCCGCAAAACCTCGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCCGACCCGCAAAACCTCGGAA 1920
Qy 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAGCGAAAGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAGCGAAAGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTTCACCTGGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCCTGGTTGACCTGATTAATAAACAACAACACATGAATGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCCTGGTTGACCTGATTAATAAACAACAACACATGAATGAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAACTGCTGCTTAATAAGGCGAAACAGCGATGACCATCAACGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAACTGCTGCTTAATAAGGCGAAACAGCGATGACCATCAACGCGCCGCTGG 2220
Qy 2221 GCATGGTCCAACTCGACACCGCAAAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAACTCGACACCGCAAAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
Qy 2281 AAGGTTCAACCATCAACCGTTGCTGGCTGCTGAGCGGAGGTATTAACGCGCCGAGT 2340
Db 2281 AAGGTTCAACCATCAACCGTTGCTGGCTGCTGAGCGGAGGTATTAACGCGCCGAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGGTTTAATAAGCAAAACCGCTGGGTGCGGTAGCGTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTTAATAAGCAAAACCGCTGGGTGCGGTAGCGTGAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGCGAAGACATCCAGTATTTGCGGCCACCATGGAACCCCGAGAAAGTGAATCATG 2520
Db 2461 TTGCGAAGACATCCAGTATTTGCGGCCACCATGGAACCCCGAGAAAGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580

Db 2521 CCGAACATCCCGCAGATGTCGGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGAAGACGCGCAGACTAAATTCGAGCTCG 2640
Qy 2641 AACCAACAACAACATTAACAATTAACAACCTCGGGATCGAGGAGAGATTTTCAGAATTC 2700
Db 2641 AACCAACAACAACATTAACAATTAACAACCTCGGGATCGAGGAGAGATTTTCAGAATTC 2700
Qy 2701 GGATCTCTTCTCTGTCGGCCAGGGCGCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTCGGCCAGGGCGCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
Qy 2761 CCGGAATGTGGTAACTCTCTCAGAGCTCTCACCTGGTGGCCACACGAGCGTACCCAC 2820
Db 2761 CCGGAATGTGGTAACTCTCTCAGAGCTCTCACCTGGTGGCCACACGAGCGTACCCAC 2820
Qy 2821 ACGGTGAAAAACCGGTATAAATGCCAGAGTGGGCAAAATCTTTTAGCACCAGGGGCTCC 2880
Db 2821 ACGGTGAAAAACCGGTATAAATGCCAGAGTGGGCAAAATCTTTTAGCACCAGGGGCTCC 2880
Qy 2881 CTGGTGGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCCAGAAATGTGGC 2940
Db 2881 CTGGTGGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCCAGAAATGTGGC 2940
Qy 2941 AAGTCTTTAGCAGCAGCTCCAGCCTGTCGCGCCACCAAGTACTCACACCGGGGAGAAG 3000
Db 2941 AAGTCTTTAGCAGCAGCTCCAGCCTGTCGCGCCACCAAGTACTCACACCGGGGAGAAG 3000
Qy 3001 CCTATGCTTGTCCGGAATGTGGTAACTCTCAGCCAGAGCAGCTCCCTGGTGGCCAC 3060
Db 3001 CCTATGCTTGTCCGGAATGTGGTAACTCTCAGCCAGAGCAGCTCCCTGGTGGCCAC 3060
Qy 3061 CAGCGTACCACACCGCGTGAATAACCGTATAATGCCAGAGTCCGCGCAAAATCTTTTAGT 3120
Db 3061 CAGCGTACCACACCGCGTGAATAACCGTATAATGCCAGAGTCCGCGCAAAATCTTTTAGT 3120
Qy 3121 GACTGCGCGGACCTTCTGCGCATCAAGCCTCATCTGCGGAGAGCCATCAAAATGT 3180
Db 3121 GACTGCGCGGACCTTCTGCGCATCAAGCCTCATCTGCGGAGAGCCATCAAAATGT 3180
Qy 3181 CCAGAAATGTGGCAAGTCTTTCTCCCAATCCAGCATCTCGTCCGCGCAACGCTACTCAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTTCTCTCTGCTGCAATCTGTCGCGCACCAAGCTACTCAC 3240
Qy 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGCCAGTACCCGTACGACGTTCCGGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGCCAGTACCCGTACGACGTTCCGGACTACGCT 3300

RESULT 4

AX202427 LOCUS 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 15 from Patent WO0152620.
ACCESSION AX202427
VERSION AX202427.1 GI:15392175
KEYWORDS .

SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.

REFERENCE

1
Barbas, C.F., Stege, J.T., Guan, X. and Dalmia, B.
Methods and compositions to modulate expression in plants
Patent: WO 0152620-A 15 26-JUL-2001;
The Scripps Research Institute (US); SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)

FEATURES
source
1. .3300
Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-m2 and zinc finger protein

ORIGIN									
ZFPm2"									
Query Match									
Best Local Similarity 97.2%; Score 3207.2; DB 6; Length 3300;									
Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;									
QY	1	CCGACACCATCGAATGGTGGCAAAACCTTTCGGGTATGGCATATAGCGCCCGGAGAGA	60						
DB	1	CCGACACCATCGAATGGTGGCAAAACCTTTCGGGTATGGCATATAGCGCCCGGAGAGA	60						
QY	61	GTCAATTCAGGGTGGTGAATGTGAACACCAAGTAAACGATATACGATGTCGACAGTATGCCG	120						
DB	61	GTCAATTCAGGGTGGTGAATGTGAACACCAAGTAAACGATATACGATGTCGACAGTATGCCG	120						
QY	121	GTGTCTTTATCAGACCGTTTCCCGCGTGGTGAACACGCGAGCCACGCTTTCGCGAAAA	180						
DB	121	GTGTCTTTATCAGACCGTTTCCCGCGTGGTGAACACGCGAGCCACGCTTTCGCGAAAA	180						
QY	181	CGCGGAAAGTGAAGCGCGGATGGCGGAGCTGAATTAATTCCTCCACCGCTGGGCAC	240						
DB	181	CGCGGAAAGTGAAGCGCGGATGGCGGAGCTGAATTAATTCCTCCACCGCTGGGCAC	240						
QY	241	AACAACTGCGGGCAAAACAGTCTGTCTGATTGGCGTTTGGCACTCCAGTCTGGCCCTGC	300						
DB	241	AACAACTGCGGGCAAAACAGTCTGTCTGATTGGCGTTTGGCACTCCAGTCTGGCCCTGC	300						
QY	301	ACGCGCGTTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGC	360						
DB	301	ACGCGCGTTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGC	360						
QY	361	TGGTGGTTCGATGTAGAAACGAGCGGCTGAAAGCGGCTGAAAGCGGCTGCAATC	420						
DB	361	TGGTGGTTCGATGTAGAAACGAGCGGCTGAAAGCGGCTGAAAGCGGCTGCAATC	420						
QY	421	TTCTCGCCCAACGGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATCAGCAGGATGCCA	480						
DB	421	TTCTCGCCCAACGGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATCAGCAGGATGCCA	480						
QY	481	TTGCTGTGGAAAGCTGCCTGCACATAATGTTTCGGCGCTTAATTTCTGATGTCTCTGACCCAGA	540						
DB	481	TTGCTGTGGAAAGCTGCCTGCACATAATGTTTCGGCGCTTAATTTCTGATGTCTCTGACCCAGA	540						
QY	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600						
DB	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600						
QY	601	TGGTGGCATTTGGTCAACAGAAATTCGGCTGTTAGCGGCGCCCATTAAGTCTGTCTCGG	660						
DB	601	TGGTGGCATTTGGTCAACAGAAATTCGGCTGTTAGCGGCGCCCATTAAGTCTGTCTCGG	660						
QY	661	CGCGTCTCGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
DB	661	CGCGTCTCGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
QY	721	CGGAAACGGAAAGCGACTGGAGTGCATATGTCGGTTTTCAACAAACCATGCAAAATGCTGA	780						
DB	721	CGGAAACGGAAAGCGACTGGAGTGCATATGTCGGTTTTCAACAAACCATGCAAAATGCTGA	780						
QY	781	ATGAGGGCATGTTCCCATCTGGATGCTGGTTCGCAACGATCAGATGGCGTGGCGCGCA	840						
DB	781	ATGAGGGCATGTTCCCATCTGGATGCTGGTTCGCAACGATCAGATGGCGTGGCGCGCA	840						
QY	841	TGCGCGCATTAACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG	900						
DB	841	TGCGCGCATTAACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG	900						
QY	901	ACGATACCGAAAGACAGCTCATGTTATATCCCGCGTTAAACCAACATCAACAGGATTTTC	960						
DB	901	ACGATACCGAAAGACAGCTCATGTTATATCCCGCGTTAAACCAACATCAACAGGATTTTC	960						
QY	961	GCCTGCTGGGCAAAACCAAGCTGGAGCCGCTTGCTGCAACTCTCTCAGGCGCAGCGGTGA	1020						
DB	961	GCCTGCTGGGCAAAACCAAGCTGGAGCCGCTTGCTGCAACTCTCTCAGGCGCAGCGGTGA	1020						

DB	961	GCCTGCTGGGCAAAACCAAGCTGGAGCCGCTTGCTGCAACTCTCTCAGGCGCAGCGGTGA	1020
QY	1021	AGGCAATCAGCTGTGGCCGCTCTCACTGTGTGAAAGAAACCAACCTTGCGGCCCAATA	1080
DB	1021	AGGCAATCAGCTGTGGCCGCTCTCACTGTGTGAAAGAAACCAACCTTGCGGCCCAATA	1080
QY	1081	CGCAACCGCCTCTCCCGCGGTTGGCCGATTCATTAATCAGCTGGCAGACAGGTTT	1140
DB	1081	CGCAACCGCCTCTCCCGCGGTTGGCCGATTCATTAATCAGCTGGCAGACAGGTTT	1140
QY	1141	CCGACTGGAAAGCGGCGAGTGAGCGCAACCGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200
DB	1141	CCGACTGGAAAGCGGCGAGTGAGCGCAACCGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200
QY	1201	GCACAAATTCATGTTTGACAGCTTATCATCGACTGCGACCGTGCACCAATGTTCTGGCG	1260
DB	1201	GCACAAATTCATGTTTGACAGCTTATCATCGACTGCGACCGTGCACCAATGTTCTGGCG	1260
QY	1261	TCAGCGAGCCATCGGAAGCTGTGTATGGCTGTGCAGTCTGTAATCACTGCATAATTCG	1320
DB	1261	TCAGCGAGCCATCGGAAGCTGTGTATGGCTGTGCAGTCTGTAATCACTGCATAATTCG	1320
QY	1321	TGTGCTCAAGCGGCACCTCCCGTTCTGGATTAATGTTTTTCGGCGGACATCATACCGGTT	1380
DB	1321	TGTGCTCAAGCGGCACCTCCCGTTCTGGATTAATGTTTTTCGGCGGACATCATACCGGTT	1380
QY	1381	CTGCAAAATTTCTGAAATGAGCTGTTGAACAATTAATCATCGGCTCGTAAATGTTGTGGA	1440
DB	1381	CTGCAAAATTTCTGAAATGAGCTGTTGAACAATTAATCATCGGCTCGTAAATGTTGTGGA	1440
QY	1441	ATTGTGAGCGGATTAACAATTTTCAACAGGAACAGCCAGTCCGTTTAGTGTTCACGA	1500
DB	1441	ATTGTGAGCGGATTAACAATTTTCAACAGGAACAGCCAGTCCGTTTAGTGTTCACGA	1500
QY	1501	GCACCTTCAACAAAGGACCATAGATTATGAAACTGAAGAAAGGTAAACTGGTAATCTGG	1560
DB	1501	GCACCTTCAACAAAGGACCATAGATTATGAAACTGAAGAAAGGTAAACTGGTAATCTGG	1560
QY	1561	ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCCGTTAGGTGTTTTCACGA	1620
DB	1561	ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCCGTTAGGTGTTTTCACGA	1620
QY	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT	1680
DB	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT	1680
QY	1681	CGCGCACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC	1740
DB	1681	CGCGCACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCCTGTTGGCTGAAATCAACCCGGAACAAAGCGTTCCAGGACAAAGCTGAT	1800
DB	1741	GCTCAATCTGGCCTGTTGGCTGAAATCAACCCGGAACAAAGCGTTCCAGGACAAAGCTGAT	1800
QY	1801	CGGTTTACCTGGGATGCGGTACGTTTAAACAGAGATCTGCTCCGAAACCCGCAACCTGGGAA	1860
DB	1801	CGGTTTACCTGGGATGCGGTACGTTTAAACAGAGATCTGCTCCGAAACCCGCAACCTGGGAA	1860
QY	1861	GAAGGTTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCAACCTGGGAA	1920
DB	1861	GAAGGTTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCAACCTGGGAA	1920
QY	1921	GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980
DB	1921	GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTTCACCTGGCGCTGATGTTGCTGACGCGGGGTTATGCGTTCAAG	2040
DB	1981	CTGCAAGAACCGTACTTTCACCTGGCGCTGATGTTGCTGACGCGGGGTTATGCGTTCAAG	2040
QY	2041	TATGAAAACCGCAAGTACGACATTAAGACCTGGCGGTGATTAACGCTGGCGCGAAAGCG	2100
DB	2041	TATGAAAACCGCAAGTACGACATTAAGACCTGGCGGTGATTAACGCTGGCGCGAAAGCG	2100

QY 2101 GGTCTGACCTTCTGTTGACCTGATTAAACAAACACATGAATGACACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTGACCTGATTAAACAAACACATGAATGACACACCGATTAC 2160
QY 2161 TCCATCGCAGAGTGCCTTTAATAAGCGGAAACAGCGATGACCATCAACCGGCCGTGG 2220
Db 2161 TCCATCGCAGAGTGCCTTTAATAAGCGGAAACAGCGATGACCATCAACCGGCCGTGG 2220
QY 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTATGGTGAACGGTACTGCCACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTATGGTGAACGGTACTGCCACCTTC 2280
QY 2281 AAGGTCACACCATCAACACCGTTCGTTGGCGTGTGAGCGAGGTATTAACCGCCGACGT 2340
Db 2281 AAGGTCACACCATCAACACCGTTCGTTGGCGTGTGAGCGAGGTATTAACCGCCGACGT 2340
QY 2341 CCGAACAAAGAGTGGGAAAAGAGTTCCTCGAAACATCTCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGTGGGAAAAGAGTTCCTCGAAACATCTCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGTCCGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTCCGTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGTTATTGCGGCCACCATGGAAAACGCCGAGAAAGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATTGCGGCCACCATGGAAAACGCCGAGAAAGTGAATCATG 2520
QY 2521 CCGAACATCCGCGAGATGTCGCTTTCTGGTATGCGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCGCGAGATGTCGCTTTCTGGTATGCGCTGCTGCTGCTGCTGCTGCTG 2580
QY 2581 GCCAGCGTCTGACAGTCTGATGAAGCCCTGGAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTCTGATGAAGCCCTGGAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAATAACAATAACAACCTCGGGATCGAGGGAAGATTCAGAAATTC 2700
Db 2641 AACAAACAAACAATAACAATAACAACCTCGGGATCGAGGGAAGATTCAGAAATTC 2700
QY 2701 GGATCTCTTCTCTGTGGCCAGCGGCCCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTGGCCAGCGGCCCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTGTGAAGTCTCTCAGCGATCTCGGCCACCTGGTTCCGCCACCGCGTACCAC 2820
Db 2761 CCGGAATGTGTGAAGTCTCTCAGAGCTCTCACCTGGTGCACCGCGTACCAC 2820
QY 2821 ACGGGTGAACACCGTATAATGCCAGAGTGGCGCAAACTTTTAGCACCGCGGCTCC 2880
Db 2821 ACGGGTGAACACCGTATAATGCCAGAGTGGCGCAAACTTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGGTGGCCATCAACGCTACTACTGGCGAGAGCGCATACAATGTCCGAATGTGGC 2940
Db 2881 CTGGTGGCCATCAACGCTACTACTGGCGAGAGCGCATACAATGTCCGAATGTGGC 2940
QY 2941 AAGTCTTTTACGCCAGAGCTCCAGCGCTGTGGCGCCACCAAGCTACTCACCGGGGAGAAG 3000
Db 2941 AAGTCTTTTCTCGTCTGACAACTCTGTTCGGGACCAACGCTACTCACCGGGGAGAAG 3000
QY 3001 CCTTATGCTTGTCCGGAATGTGGTAAGTCTTTAGCCAGAGCAGCTCCCTGGTGGCGCAC 3060
Db 3001 CCTTATGCTTGTTCGGAATGTGGTAAGTCTTTAGCCGCGCAGATACCTGGTGGCGCAC 3060
QY 3061 CAGGTTACCCACACGGGTGAACACCGTATAAATGCCGAGAGTGGCGCAATCTTTTAGT 3120
Db 3061 CAGGTTACCCACACGGGTGAACACCGTATAAATGCCGAGAGTGGCGCAATCTTTTAGC 3120
QY 3121 GACTGCGCGCACCTTGTCTCGGCATCAACGCACTCATACTGCGGAGAGGCCATACAATGT 3180
Db 3121 CAGCGCGCACCTTGGCCAGGCCATCAACGCACTCATACTGCGGAGAGGCCATACAATGT 3180

QY 3181 CCAGATGTGGCAAGTCTTTCTCCCAATCCAGCCATCTGTCGGGACCAACGTTACTCAC 3240
Db 3181 CCAGATGTGGCAAGTCTTTCTCTCGTCTGACAAATCTCGTCCGACCAACGTTACTCAC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGGCGGACCGTACCGAGTTCGGGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGGCGGACCGTACCGTACGAGTTCGGGACTACGCT 3300
RESULT 5
AX202429 3300 bp DNA linear PAT 30-AUG-2001
LOCUS Sequence 17 from Patent WO0152620.
DEFINITION AX202429
ACCESSION AX202429
VERSION AX202429.1 GI:153921177
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 other sequences; artificial sequences.
AUTHORS Barbas, C.F., Stege, J.T., Guan, X. and Dalmia, B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 17 26-JUL-2001;
The Scripps Research Institute (US); SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES
Location/Qualifiers
source 1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-m4 and zinc finger protein ZFPm4"

ORIGIN
Query Match 97.2%; Score 3207.2; DB 6; Length 3300;
Beet Local Similarity 98.2%; Pred. No. 0;
Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 1 CCACACCATCGAATGGTGCAAAACCTTTTCGGGTATGGCATATAGGCCCGGAAGAGA 60
Db 1 CCACACCATCGAATGGTGCAAAACCTTTTCGGGTATGGCATATAGGCCCGGAAGAGA 60
QY 61 GTCAATTCAGGGTGGTGAATGTGAAACCACTAATACGATGTGCGAGAGTATGCCG 120
Db 61 GTCAATTCAGGGTGGTGAATGTGAAACCACTAATACGATGTGCGAGAGTATGCCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACAGAGCCAGCCAGCTTTCTGCGAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACAGAGCCAGCCAGCTTTCTGCGAAA 180
QY 181 CCGGGGAAAAGTGGAGCGCGCATGGCGAGCTGAATTAACATTCCCAACCGGTTGGCAC 240
Db 181 CCGGGGAAAAGTGGAGCGCGCATGGCGAGCTGAATTAACATTCCCAACCGGTTGGCAC 240
QY 241 AACCACTGGCGGGCAACAGTCTGCTGATTGGCGTTGGCACCTCCAGTCTGGCCCTGC 300
Db 241 AACCACTGGCGGGCAACAGTCTGCTGATTGGCGTTGGCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTCCGAAATTTGTGCGCGCGATTAATCTCGCGCGGATCACTGGGTGCCAGCG 360
Db 301 ACGCGCGTCCGAAATTTGTGCGCGCGATTAATCTCGCGCGGATCACTGGGTGCCAGCG 360
QY 361 TGTGTGTGTCGATGGTAGAACGAAGCGGCTGCAAGCCTGTAAGCGGCGGTGCACAATC 420
Db 361 TGTGTGTGTCGATGGTAGAACGAAGCGGCTGCAAGCCTGTAAGCGGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAACATATCCGCTGGATGACCAGGTGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAACATATCCGCTGGATGACCAGGTGCCA 480
QY 481 TTGCTGTGGAGCTGCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTTCTGACACAGA 540
Db 481 TTGCTGTGGAGCTGCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTTCTGACACAGA 540

QY	541	CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGGACTGGGCGTGGAGCATC	600
DB	541	CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGGACTGGGCGTGGAGCATC	600
QY	601	TGGTCGCATTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAAGTCTGTCTCGG	660
DB	601	TGGTCGCATTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAAGTCTGTCTCGG	660
QY	661	CGCGTCTGCGTCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	661	CGCGTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAACGGGAAGCGACTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA	780
DB	721	CGGAACGGGAAGCGACTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA	780
QY	781	ATGAGGGCATCGTCCCACTCGGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA	840
DB	781	ATGAGGGCATCGTCCCACTCGGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA	840
QY	841	TGCGGCGCAATACCGAGTCCGGGCTGCGCGTTGGTGGGATATCTCGGTAGTGGGATAG	900
DB	841	TGCGGCGCAATACCGAGTCCGGGCTGCGCGTTGGTGGGATATCTCGGTAGTGGGATAG	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCCATCAAAAGGATTTTC	960
DB	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCCATCAAAAGGATTTTC	960
QY	961	GCCTGCTGGGGCAACACGCGTGGACCGCTTCTGCACTCTCTCAGGCGCCAGGCGGTGA	1020
DB	961	GCCTGCTGGGGCAACACGCGTGGACCGCTTCTGCACTCTCTCAGGCGCCAGGCGGTGA	1020
QY	1021	AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACACCCCTGGCGGCCCAATA	1080
DB	1021	AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACACCCCTGGCGGCCCAATA	1080
QY	1081	CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGACGTGGCACACAGGTTT	1140
DB	1081	CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGACGTGGCACACAGGTTT	1140
QY	1141	CCCGACTGGAAAGCGGCGAGTACGCGCAACGAATTAATGTCAGTTAGCTCACTCATTAG	1200
DB	1141	CCCGACTGGAAAGCGGCGAGTACGCGCAACGAATTAATGTCAGTTAGCTCACTCATTAG	1200
QY	1201	GCACAATCTCATGTTTGAACGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG	1260
DB	1201	GCACAATCTCATGTTTGAACGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG	1260
QY	1261	TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGACGTCGTAATCACTGCATTAATTCG	1320
DB	1261	TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGACGTCGTAATCACTGCATTAATTCG	1320
QY	1321	TGTCGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCCGACATCAACGGTT	1380
DB	1321	TGTCGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCCGACATCAACGGTT	1380
QY	1381	CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTTGGA	1440
DB	1381	CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTTGGA	1440
QY	1441	ATTGTGAGCGGATAACAATTTCAACAGAAACAGCCAGTCCGTTAGGTGTTTCAACA	1500
DB	1441	ATTGTGAGCGGATAACAATTTCAACAGAAACAGCCAGTCCGTTAGGTGTTTCAACA	1500
QY	1501	GCACCTCACCAACAGGACCATAGATTATGAACCTGAAGAGGTAAACTGTAATCTGG	1560
DB	1501	GCACCTCACCAACAGGACCATAGATTATGAACCTGAAGAGGTAAACTGTAATCTGG	1560
QY	1561	ATTAAACGGGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAGAAAT	1620
DB	1561	ATTAAACGGGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAGAAAT	1620

QY	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAGAGAAATTTCCACAGGTT	1680
DB	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAGAGAAATTTCCACAGGTT	1680
QY	1681	GGGCAACTGCGCATGGGCGCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC	1740
DB	1681	GGGCAACTGCGCATGGGCGCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGCGCTGTTGGCTGAAATCAACCCCGCAAAAGCGTTCCAGGACAAGCTGTAT	1800
DB	1741	GCTCAATCTGCGCTGTTGGCTGAAATCAACCCCGCAAAAGCGTTCCAGGACAAGCTGTAT	1800
QY	1801	CCGTTTACCTGGGATGCGGTACGTTACAAACGCAAGCTGATTGCTTACCCGATCCCTGTT	1860
DB	1801	CCGTTTACCTGGGATGCGGTACGTTACAAACGCAAGCTGATTGCTTACCCGATCCCTGTT	1860
QY	1861	GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA	1920
DB	1861	GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA	1920
QY	1921	GAGATCCCGGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
DB	1921	GAGATCCCGGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGGGTTATCGGTTCAAG	2040
DB	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGGGTTATCGGTTCAAG	2040
QY	2041	TATGAAAAACGGCAAGTACGACATTAAGACGTTGGCGCTGGATAACGCTGGCGGAAAGCG	2100
DB	2041	TATGAAAAACGGCAAGTACGACATTAAGACGTTGGCGCTGGATAACGCTGGCGGAAAGCG	2100
QY	2101	GCTCTGACCTTCTCGTTGAACTGATTAATAAACAACATGAAATGCGAGACACCGATTAC	2160
DB	2101	GCTCTGACCTTCTCGTTGAACTGATTAATAAACAACATGAAATGCGAGACACCGATTAC	2160
QY	2161	TCCATCGCAGAGCTGCGCTTTAATAAGCGGAAACAGCGATGACCATCAACCGCGCTGG	2220
DB	2161	TCCATCGCAGAGCTGCGCTTTAATAAGCGGAAACAGCGATGACCATCAACCGCGCTGG	2220
QY	2221	GCATGGTCAACATCGACACAGCAAAAGTGAATTAATGTTAAACGCTACTGCGGACCTTC	2280
DB	2221	GCATGGTCAACATCGACACAGCAAAAGTGAATTAATGTTAAACGCTACTGCGGACCTTC	2280
QY	2281	AAGGTTCAACCATCAAAACCGTTCTGTCGCTGAGCGCAGGTAATTAACCGCCCGACT	2340
DB	2281	AAGGTTCAACCATCAAAACCGTTCTGTCGCTGAGCGCAGGTAATTAACCGCCCGACT	2340
QY	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAANAATATCTGCTGACTGATGAAGTCTG	2400
DB	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAANAATATCTGCTGACTGATGAAGTCTG	2400
QY	2401	GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAG	2460
DB	2401	GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAG	2460
QY	2461	TTGGCGAAAGATCCACGTAATTTGCCGCCACCATGGAAGAAACGCCAGAAAGTGAAATCATG	2520
DB	2461	TTGGCGAAAGATCCACGTAATTTGCCGCCACCATGGAAGAAACGCCAGAAAGTGAAATCATG	2520
QY	2521	CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTGCTGCTGATCAACGCG	2580
DB	2521	CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTGCTGCTGATCAACGCG	2580
QY	2581	GCCAGCGGTGCTGACATGTCGATGAAGCGCTGAAAGAGCGCGACACTAATTCGAGCTCG	2640
DB	2581	GCCAGCGGTGCTGACATGTCGATGAAGCGCTGAAAGAGCGCGACACTAATTCGAGCTCG	2640
QY	2641	AACAAACAACAATAACAATAACCAACCTCGGGATCGAGGGAAGGATTTTCAAGATTC	2700
DB	2641	AACAAACAACAATAACAATAACCAACCTCGGGATCGAGGGAAGGATTTTCAAGATTC	2700
QY	2701	GGATCCTCTTCTCTGCTGGCCCGAGGCGGCTTCGAGACCCCGGGGAGAACCCCTATGCTGT	2760


```
QY 841 TCGCGGCCATTACCGAGTCCGGGCTCGCGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TCGCGGCCATTACCGAGTCCGGGCTCGCGTTGGTGGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTGTATATCCCGCGTTTAACCAACCATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTGTATATCCCGCGTTTAACCAACCATCAAAACAGGATTTTC 960
QY 961 GCCTCTGGGGCAAAACCGAGCTGGACCGGTTGCTGCAACTCTCTCAGGGCCACGGCGGTGA 1020
Db 961 GCCTCTGGGGCAAAACCGAGCTGGACCGGTTGCTGCAACTCTCTCAGGGCCACGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGGCCGTTCTCATCTGGTGAAGAAAAACCAACCTCTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCGTTCTCATCTGGTGAAGAAAAACCAACCTCTGGCGGCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGATTCAATTAATGAGCTGGCGACCAAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGATTCAATTAATGAGCTGGCGACCAAGGTTT 1140
QY 1141 CCCGACTGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCAATTAG 1200
Db 1141 CCCGACTGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCAATTAG 1200
QY 1201 GCACAATTCTCATGTTTGACAGCTTATCATCGACTGCGAGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTCTCATGTTTGACAGCTTATCATCGACTGCGAGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGCGAGCATCGGAAGCTGTGGTATGCTGTGAGGTGTAATCACTGCGCATTAATTCG 1320
Db 1261 TCAGGCGAGCATCGGAAGCTGTGGTATGCTGTGAGGTGTAATCACTGCGCATTAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTTGGCGGACATCAATAACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTTGGCGGACATCAATAACGGTT 1380
QY 1381 CTGGCAATATTTCTGAATAGCTGTGTGACAATTAATCATCGGCTCGTATAATGTGTGA 1440
Db 1381 CTGGCAATATTTCTGAATAGCTGTGTGACAATTAATCATCGGCTCGTATAATGTGTGA 1440
QY 1441 ATTGTGAGCGGATACAAATTTACACAGGAACACCGCTGCTGAGGTGTAATCACTGCGCATTAATTCG 1500
Db 1441 ATTGTGAGCGGATACAAATTTACACAGGAACACCGCTGCTGAGGTGTAATCACTGCGCATTAATTCG 1500
QY 1501 GCACCTTCAACAAAGGACCATAGATTATGAANAATCGAAGAGGTAACTGCTAATCTCG 1560
Db 1501 GCACCTTCAACAAAGGACCATAGATTATGAANAATCGAAGAGGTAACTGCTAATCTCG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACGGTTCAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGGTTCAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCAATGCGGATGGCCCTGACATTATCTTGGGACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAATGCGGATGGCCCTGACATTATCTTGGGACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGGCTGAATTCACCCCGGACAAAGCGTTTCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAATTCACCCCGGACAAAGCGTTTCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCGGTAGCTTACACGCGAAGCTGATGCTTACCGGATCGCTGT 1860
Db 1801 CCGTTTACCTGGATGCGGTAGCTTACACGCGAAGCTGATGCTTACCGGATCGCTGT 1860
QY 1861 GAAGCGTTATCGTGAATTTATACAAAGATCTGTCGCAACCCGCGCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGTGAATTTATACAAAGATCTGTCGCAACCCGCGCAAAACCTGGGAA 1920
```

```
QY 1921 GAGATCCCGGCTGGATAAAGNACTGAAGCGGAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCTGGATAAAGNACTGAAGCGGAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCCACTGGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCCACTGGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGAACGTTGGCGCTGGATAAACGTTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGAACGTTGGCGCTGGATAAACGTTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCCCTGGTTGACCTGATTAATAAACACATCAATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCCCTGGTTGACCTGATTAATAAACACATCAATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCCCTTTAATAAAGCGAAACAGCGATGACCATCAACGGGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCCCTTTAATAAAGCGAAACAGCGATGACCATCAACGGGCCGCTGG 2220
QY 2221 GCATGGTCCAAACATCGACACCAAGTAAATTAATGTGTAAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAACATCGACACCAAGTAAATTAATGTGTAAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGTCACACCATCCAAACCGTTGCTGGCGTCTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTCACACCATCCAAACCGTTGCTGGCGTCTGAGCGCAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCAAGATCCACGTTATTCGCCCAACCATGGAAGAACCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCAAGATCCACGTTATTCGCCCAACCATGGAAGAACCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCTGCTGCTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGGAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGGAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACACACACAAATAACAAATAACAAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACACACACAAATAACAAATAACAAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
QY 2701 GGATCCTCTTCTCTGTGGCCCGCGCGCTCTGAGCGCGG 2741
Db 2701 GGATCCTCTGATCCCGCAAGCCCGCGCTCTGTTGGGAGG 2741

RESULT 7
AX172306
LOCUS AX172306 8101 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 5 from Patent WO014467.
ACCESSION AX172306
VERSION AX172306.1 GI:14597487
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGNISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Goubin-Gramatica,F., Ducommun,B. and Prevost,G.
TITILE Method for obtaining human cdc25 phosphatases and method for
identifying human cdc25 phosphatase modulators
JOURNAL Patent: WO 014467-A 5 21-JUN-2001;
SOCIETE DE CONSEILS DE RECHERCHES ET D'APPLICATIONS SCIENTIFIQUES
```


FEATURES		(FR)		Location/Qualifiers	
source		1..8101		/organism="synthetic construct"	
				/mol_type="unassigned DNA"	
				/db_xref="taxon:32630"	
				/note="souche JM109 / pMAL - Hs Cdc25C"	
ORIGIN					
Query Match 82.0%; Score 2705.8; DB 6; Length 8101;					
Best Local Similarity 99.9%; Pred. No. 0;					
Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;					
Qy	1	CCGACACCATCGAATGGTGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA	60		
Db	1	CCGACACCATCGAATGGTGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA	60		
Qy	61	GTCAATTTCAGGGTGGTGAATGTGAACCAAGTAAGCTTATACGATGTGCGAGATATGCGG	120		
Db	61	GTCAATTTCAGGGTGGTGAATGTGAACCAAGTAAGCTTATACGATGTGCGAGATATGCGG	120		
Qy	121	GTGTCTCTTATCAGACCGGTTTCCCGGTGGTGAACCGCCAGCCAGCTTTCTCGGAAA	180		
Db	121	GTGTCTCTTATCAGACCGGTTTCCCGGTGGTGAACCGCCAGCCAGCTTTCTCGGAAA	180		
Qy	181	CGCGGGAAGAGTGGAGCGCGCATGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC	240		
Db	181	CGCGGGAAGAGTGGAGCGCGCATGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC	240		
Qy	241	AACAATCGCGGGCAACAGTCGTTGCTGATGGGGTGGCCACTCCAGTCTGCGCCCTGC	300		
Db	241	AACAATCGCGGGCAACAGTCGTTGCTGATGGGGTGGCCACTCCAGTCTGCGCCCTGC	300		
Qy	301	ACGCGCGTTCGCAAAATTTGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360		
Db	301	ACGCGCGTTCGCAAAATTTGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360		
Qy	361	TGTTGGTTCGATGGTGAACGAAGCGCGCTGGAAGCTGTGAAGCGCGGTGCACAATC	420		
Db	361	TGTTGGTTCGATGGTGAACGAAGCGCGCTGGAAGCTGTGAAGCGCGGTGCACAATC	420		
Qy	421	TTCTCGCGCAACCGTCAAGTGGGCTGATCAATTAATCTCGCGCGATCAACTGGGTGCCAG	480		
Db	421	TTCTCGCGCAACCGTCAAGTGGGCTGATCAATTAATCTCGCGCGATCAACTGGGTGCCAG	480		
Qy	481	TTGCTGTGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTTGACCGA	540		
Db	481	TTGCTGTGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTTGACCGA	540		
Qy	541	CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGGACTGGCGGTGGAGATC	600		
Db	541	CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGGACTGGCGGTGGAGATC	600		
Qy	601	TGGTTCGCAATTTGGGTCAACAGCAATCGCTGTTAGCGGGCCCAATTAAAGTTCTGTCGG	660		
Db	601	TGGTTCGCAATTTGGGTCAACAGCAATCGCTGTTAGCGGGCCCAATTAAAGTTCTGTCGG	660		
Qy	661	CGGCTCTGGCTGCGTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720		
Db	661	CGGCTCTGGCTGCGTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720		
Qy	721	CGGAACGGGAAGGCACTGGAGTCCCATGTCGGTTTCAACAAACCATGCAATGTCTGA	780		
Db	721	CGGAACGGGAAGGCACTGGAGTCCCATGTCGGTTTCAACAAACCATGCAATGTCTGA	780		
Qy	781	ATGAGGCACTGTTCCCACTCGCATGCTGGTTGGCAACAGATCAGATGGCGCTGGCGCAA	840		
Db	781	ATGAGGCACTGTTCCCACTCGCATGCTGGTTGGCAACAGATCAGATGGCGCTGGCGCAA	840		
Qy	841	TGCGCGCATTAACGAGTCCGGGTGCGGTGGTGGGATATCTCGGTAGTGGGATACG	900		
Db	841	TGCGCGCATTAACGAGTCCGGGTGCGGTGGTGGGATATCTCGGTAGTGGGATACG	900		

Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGCTTAACACCATCAAAACAGGATTTTC	960		
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGCTTAACACCATCAAAACAGGATTTTC	960		
Qy	961	GCTGTCTGGGGCAAAACAGCGTGGACCGCTTGTGTCAACTCTCTCAGGCGCAGCGGTGA	1020		
Db	961	GCTGTCTGGGGCAAAACAGCGTGGACCGCTTGTGTCAACTCTCTCAGGCGCAGCGGTGA	1020		
Qy	1021	AGGCAATCAGCTGTGGCGCTCTCACTGTGAAAGAAAACACCTGGCGCCCAATA	1080		
Db	1021	AGGCAATCAGCTGTGGCGCTCTCACTGTGAAAGAAAACACCTGGCGCCCAATA	1080		
Qy	1081	CGAAAACCGCTCTCCCGCGGCTTGGCCGATTAATTAATGCAGCTGGCAGCAGGTTT	1140		
Db	1081	CGAAAACCGCTCTCCCGCGGCTTGGCCGATTAATTAATGCAGCTGGCAGCAGGTTT	1140		
Qy	1141	CCGCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGATTAAGCTCACTCATTTAG	1200		
Db	1141	CCGCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGATTAAGCTCACTCATTTAG	1200		
Qy	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG	1260		
Db	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG	1260		
Qy	1261	TCAGGCAAGCCATCGGAAGCTGTGTATGGCTGTGCAGCTCGTAATCACTGCATAATTCG	1320		
Db	1261	TCAGGCAAGCCATCGGAAGCTGTGTATGGCTGTGCAGCTCGTAATCACTGCATAATTCG	1320		
Qy	1321	TGTGCTCAAGGCGCACTCCGTTCTGGATAATGTTTTTGGCGCGGACATCATAAACGGTT	1380		
Db	1321	TGTGCTCAAGGCGCACTCCGTTCTGGATAATGTTTTTGGCGCGGACATCATAAACGGTT	1380		
Qy	1381	CTGCGAAATATCTGAAATGAGCTGTTGCAAAATTAATCATCGCTCGTATAATGTGTGGA	1440		
Db	1381	CTGCGAAATATCTGAAATGAGCTGTTGCAAAATTAATCATCGCTCGTATAATGTGTGGA	1440		
Qy	1441	ATTGTGAGCGGATAACAAATTTCAACAGGAAACAGCCAGTCCGTTTAGTGTGTTTTCACGA	1500		
Db	1441	ATTGTGAGCGGATAACAAATTTCAACAGGAAACAGCCAGTCCGTTTAGTGTGTTTTCACGA	1500		
Qy	1501	GCACTTCAACCAACAGGACCATAGATTATGAAAATCGAAGAGGTAAACTGGTAACTCGG	1560		
Db	1501	GCACTTCAACCAACAGGACCATAGATTATGAAAATCGAAGAGGTAAACTGGTAACTCGG	1560		
Qy	1561	ATTAAACGGCGATTAAGGCTATAACGGTCTCGCTGAAGTGGTAAGAAATTCGAGAAAGAT	1620		
Db	1561	ATTAAACGGCGATTAAGGCTATAACGGTCTCGCTGAAGTGGTAAGAAATTCGAGAAAGAT	1620		
Qy	1621	ACCGAAATTAAGTCAACCGTTGAGCATTAACCGGATAAATCGGAAGAGAAATTCACAGGTT	1680		
Db	1621	ACCGAAATTAAGTCAACCGTTGAGCATTAACCGGATAAATCGGAAGAGAAATTCACAGGTT	1680		
Qy	1681	CGCGCAACTGGCGATGGCGCTGACATTAATCTGGGCAACAGCCGCTTGGTGGCTAC	1740		
Db	1681	CGCGCAACTGGCGATGGCGCTGACATTAATCTGGGCAACAGCCGCTTGGTGGCTAC	1740		
Qy	1741	GCTCAATCTGGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT	1800		
Db	1741	GCTCAATCTGGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT	1800		
Qy	1801	CCGTTTACCTGGGATGCGGTACGTTACACGGCAAGCTGATTTCCGATCGCTGTTT	1860		
Db	1801	CCGTTTACCTGGGATGCGGTACGTTACACGGCAAGCTGATTTCCGATCGCTGTTT	1860		
Qy	1861	GAAGCGTTATCGCTGATTTATACAAAGATCTGCTCCGAAACCCGCCCAAAACCTGGGAA	1920		
Db	1861	GAAGCGTTATCGCTGATTTATACAAAGATCTGCTCCGAAACCCGCCCAAAACCTGGGAA	1920		
Qy	1921	GAGATCCCGCGCTGGATTAAGAACTGAAAGCGGAAAGGTAAAGACGCGCTGTGTTCAAC	1980		
Db	1921	GAGATCCCGCGCTGGATTAAGAACTGAAAGCGGAAAGGTAAAGACGCGCTGTGTTCAAC	1980		
Qy	1981	CTGCAAGAACCGTACTTTCACCTGGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG	2040		

```
Db 1981 CTGCAAGAACGCTACTTACCTGGCGCTGATTGCTGACGGGGTATTGCGTTCAAG 2040
Qy 2041 TATGAAAAACGCAAGTACGACATTAAGAACGTTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGAACGTTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAGAACGTTGAAACACATGAATGCGAGACACCATAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAGAACGTTGAAACACATGAATGCGAGACACCATAC 2160
Qy 2161 TCCATCGCAGAGCTGCTCTTAATAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTCTTAATAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
Qy 2221 GCATGGTCCAAACATCGACACCAAGAAAGTGAATATGTTGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAACATCGACACCAAGAAAGTGAATATGTTGTAACGGTACTGCGGACCTTC 2280
Qy 2281 AAGGTCAACCATCAACACCGTTCTGTTGGCTGCTGAGGCGAGGTATTAAACGCCGCCAGT 2340
Db 2281 AAGGTCAACCATCAACACCGTTCTGTTGGCTGCTGAGGCGAGGTATTAAACGCCGCCAGT 2340
Qy 2341 CGAACAAAGAGCTGCGAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CGAACAAAGAGCTGCGAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAACTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAACTTACGAGGAAGAG 2460
Qy 2461 TTGGGAAAGATCCACGATATTCGCCGCCACATGGAAGAACGCCAGAGGTGAATCATG 2520
Db 2461 TTGGGAAAGATCCACGATATTCGCCGCCACATGGAAGAACGCCAGAGGTGAATCATG 2520
Qy 2521 CGAACATCCCGCAGATGTCGCTTTCTGTTATGCGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CGAACATCCCGCAGATGTCGCTTTCTGTTATGCGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Qy 2581 GCCAGCGTCTGACAGTCTGATGAAGCCCTGAAAGAGCGCGACAGTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTCTGATGAAGCCCTGAAAGAGCGCGACAGTAAATTCGAGCTCG 2640
Qy 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Qy 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709

RESULT 8
AX377531
LOCUS AX377531 6648 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 8 from Patent WO0212553.
ACCESSION AX377531
VERSION AX377531.1 GI:19573717
KEYWORDS
SOURCE
ORGANISM
synthetic construct
synthetic construct
other sequences; artificial sequences.
REFERENCE
1
AUTHORS Kappel, A., Polakowski, T., Pignot, M., Windhab, N., Behrens, H., and Muth, J.
TITLE Method for detecting mutations in nucleotide sequences
JOURNAL Patent: WO 0212553-A 8 14-FEB-2002;
Nanogen Recognomics GmbH (DE)
FEATURES
Location/Qualifiers
1..6648
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
source
```

```
ORIGIN
Query Match 81.8%; Score 2701; DB 6; Length 6648;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2704; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
/note="Plasmid pMAL-c2x"
Qy 1 CGGACACCATGAATGGTGGAAACCTTTCCGGGTATGGCATGATAGCGCCGGAGAGA 60
Db 1 CGGACACCATGAATGGTGGAAACCTTTCCGGGTATGGCATGATAGCGCCGGAGAGA 60
Qy 61 GTCAATTTCAGGGTGGTGAATGTGAAACCAAGTAACGTTATACGATGTCGACAGATGCGG 120
Db 61 GTCAATTTCAGGGTGGTGAATGTGAAACCAAGTAACGTTATACGATGTCGACAGATGCGG 120
Qy 121 GTGCTCTTATCAGACCGTTTCCCGCGTGTGAACACGAGCCAGCCACGTTTCTGCAAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCCGCGTGTGAACACGAGCCAGCCACGTTTCTGCAAAA 180
Qy 181 CGCGGAAAAAGTGGAGCGGATGGCGGAGCTGAATATATCCCAACCGCTGGCAC 240
Db 181 CGCGGAAAAAGTGGAGCGGATGGCGGAGCTGAATATATCCCAACCGCTGGCAC 240
Qy 241 AACAACTGCGCGGAAACAGTCTGTTGCGTGTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGCGCGGAAACAGTCTGTTGCGTGTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGGCCCTGCGCAAAATTTGTCGCGCGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
Db 301 ACGGCCCTGCGCAAAATTTGTCGCGCGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
Qy 361 TGGTGGTTCGATGGTGAAGCAAGAGCGCGTGAAGCCCTGTAAGCGCGGTGCAATATC 420
Db 361 TGGTGGTTCGATGGTGAAGCAAGAGCGCGTGAAGCCCTGTAAGCGCGGTGCAATATC 420
Qy 421 TTCTCGCGCAACGCTGAGTGGGCTGATTAATTAATCTCGCTGATGACACAGGATGCCA 480
Db 421 TTCTCGCGCAACGCTGAGTGGGCTGATTAATTAATCTCGCTGATGACACAGGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGCTTATTTCTTGATGCTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGCTTATTTCTTGATGCTCTGACCCAGA 540
Qy 541 CACCCATCAACAGTATTAATTTTCCCATGAAGACGCTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTTCCCATGAAGACGCTACGCGACTGGGCGTGGAGCATC 600
Qy 601 TGGTGGATTTGGTCAACGCAAAATCGGCTGTTAGCGGGCCCATTAAGTCTGCTCGG 660
Db 601 TGGTGGATTTGGTCAACGCAAAATCGGCTGTTAGCGGGCCCATTAAGTCTGCTCGG 660
Qy 661 CGCGTCTCGCTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAAACGGAAGCGGAGCTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAAACGGAAGCGGAGCTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGCGATGTTCCCACTGGATGCTGTTGCCAAGATGAGTGGCGCTGGGCGGCA 840
Db 781 ATGAGGCGATGTTCCCACTGGATGCTGTTGCCAAGATGAGTGGCGCTGGGCGGCA 840
Qy 841 TGGCGCCATTTACCGAGTCCGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATAGC 900
Db 841 TGGCGCCATTTACCGAGTCCGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATAGC 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTGTAAACCAACCATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTGTAAACCAACCATCAAAACAGGATTTTC 960
Qy 961 GCCTGCTGGGGCAACACGAGCTGGAGCCGCTTGCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
Db 961 GCCTGCTGGGGCAACACGAGCTGGAGCCGCTTGCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
```


|||||
1 CCGACCAATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
61 GTCAATTCAGGGTGTGAATGTGAACACCAAGTAACTTATACGATGTGCGAGAGTATGCCG 120
61 GTCAATTCAGGGTGTGAATGTGAACACCAAGTAACTTATACGATGTGCGAGAGTATGCCG 120
121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCACGTTTCTCGGAAAA 180
121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCACGTTTCTCGGAAAA 180
181 CGCGGAAAAAGTGAAGCGGATGGCGGAGCTCAATTAACATCCCAACCGGTGGGCAC 240
181 CGCGGAAAAAGTGAAGCGGATGGCGGAGCTCAATTAACATCCCAACCGGTGGGCAC 240
241 AACAACTGGCGGCAAAACAGTCTGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
241 AACAACTGGCGGCAAAACAGTCTGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
301 AGCGCCGTTCGCAAAATTTGTGGCGGATTAATCTCGCGCCGATCAACTGGGTGCGACGG 360
301 AGCGCCGTTCGCAAAATTTGTGGCGGATTAATCTCGCGCCGATCAACTGGGTGCGACGG 360
361 TGGTGGTGTGATGGTGAACGAGCGGCTGAAAGCCCTGTAAAGCGGCGGTGCAACAATC 420
361 TGGTGGTGTGATGGTGAACGAGCGGCTGAAAGCCCTGTAAAGCGGCGGTGCAACAATC 420
421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACCAAGTGCAC 480
421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACCAAGTGCAC 480
481 TTGCTGTGGAAGCTCGCTGACATAATGTTCCGGCGTTATTTCTTGATCTCTGACCCAGA 540
481 TTGCTGTGGAAGCTCGCTGACATAATGTTCCGGCGTTATTTCTTGATCTCTGACCCAGA 540
541 CACCCATCAACAGTATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
541 CACCCATCAACAGTATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
601 TGGTCCGATTTGGTCCACAGCAAACTCGGCTGTTAGCGGGCCCAATTAAGTTCGTCTCGG 660
601 TGGTCCGATTTGGTCCACAGCAAACTCGGCTGTTAGCGGGCCCAATTAAGTTCGTCTCGG 660
661 CGCGTCTGCTGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
661 CGCGTCTGCTGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
721 CGGAACGGGAGGCGACCTGGAGTGCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
721 CGGAACGGGAGGCGACCTGGAGTGCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
781 ATGAGGGCATCGTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
781 ATGAGGGCATCGTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
841 TCGCGCCATTAACGAGTCCGGCTCGCGTTGGTGGGATATCTCGTGGGATAG 900
841 TCGCGCCATTAACGAGTCCGGCTCGCGTTGGTGGGATATCTCGTGGGATAG 900
901 ACGATACCGAGACAGCTCATGTTATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
901 ACGATACCGAGACAGCTCATGTTATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
961 GCCTGCTGGGCAAAACACGAGTGGACCTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
961 GCCTGCTGGGCAAAACACGAGTGGACCTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
1021 AGGCAATCAGCTGTGGCCCGTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
1021 AGGCAATCAGCTGTGGCCCGTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
1081 CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAGCAGAGTTTT 1140
|||||

Db 1081 CGCAACCGCCTCTCCCGCGGTTGGCCGATTTCATTAATGACGTGCGCAGCAGAGTTTT 1140
QY 1141 CCGCATGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCGCACTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAACGCTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAACGCTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACGCAATCGAAAGCTGTGGTGTGCTGCTGAGGTCTGTAATCACTGCTGATATTCG 1320
Db 1261 TCAGGACGCAATCGAAAGCTGTGGTGTGCTGCTGAGGTCTGTAATCACTGCTGATATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTTCGCGGCAATCAATAACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTTCGCGGCAATCAATAACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGTGGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGTGGA 1440
QY 1441 ATTGTGAGCGGATTAACAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTACAGA 1500
Db 1441 ATTGTGAGCGGATTAACAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTACAGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGGTAAAACTGGTAATCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGGTAAAACTGGTAATCTGG 1560
QY 1561 ATTTAACGGCGATAAAGGCTATAACGGTCTCGCTGAGTCCGTTAGAAATTCGAGAAAGAT 1620
Db 1561 ATTTAACGGCGATAAAGGCTATAACGGTCTCGCTGAGTCCGTTAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTTAAAGTCACCGTTGAGCATCCGATATACTGGAAGAGAAAAATCCCAAGGTT 1680
Db 1621 ACCGGAATTTAAAGTCACCGTTGAGCATCCGATATACTGGAAGAGAAAAATCCCAAGGTT 1680
QY 1681 CGCGCACTGCGCATGCGCCCTGACATATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGCGCACTGCGCATGCGCCCTGACATATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCGCAAGCTGATTTGCTTACCCGATCCGTTT 1800
Db 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCGCAAGCTGATTTGCTTACCCGATCCGTTT 1800
QY 1801 CCGTTTACCTGCGATGCGGTACGTTTACAAACGCAAGCTGATTTGCTTACCCGATCCGTTT 1860
Db 1801 CCGTTTACCTGCGATGCGGTACGTTTACAAACGCAAGCTGATTTGCTTACCCGATCCGTTT 1860
QY 1861 GAAGCGTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAACCCGCCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAACCCGCCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGGAAAGGTAAAGGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGGAAAGGTAAAGGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAAACCGTACTTCCACTGCGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAAACCGTACTTCCACTGCGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGAACGCTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGAACGCTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTTAAAAACAAACATGAAATGCAAGACCCGATTTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTTAAAAACAAACATGAAATGCAAGACCCGATTTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGGAAACAGCGATGACCATCAAGCGGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGGAAACAGCGATGACCATCAAGCGGCCGCTGG 2220

QY	2221	GCATGGTCCAACATCGACACAGCAAAAGTGAATTATGTTGTAAACGGTACTCCGACCTTC	2280
Db	2221	GCATGGTCCAACATCGACACAGCAAAAGTGAATTATGTTGTAAACGGTACTCCGACCTTC	2280
QY	2281	AAGGTCACACCATCAAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAAACGCGCCAGT	2340
Db	2281	AAGGTCACACCATCAAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAAACGCGCCAGT	2340
QY	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACCTATCTGCTGACTGATGAAGTCTG	2400
Db	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACCTATCTGCTGACTGATGAAGTCTG	2400
QY	2401	GAAGCGGTTAATAAGACAAACCGTCGGTGCCTAGCGTGAAGTCTTACGAGGAAGAG	2460
Db	2401	GAAGCGGTTAATAAGACAAACCGTCGGTGCCTAGCGTGAAGTCTTACGAGGAAGAG	2460
QY	2461	TTGCGAAAGATCCACGCTATTTCGCGCCACCATGGAACACGCCCAAGAGTGAATCATG	2520
Db	2461	TTGCGAAAGATCCACGCTATTTCGCGCCACCATGGAACACGCCCAAGAGTGAATCATG	2520
QY	2521	CCGAACATCCCGCAGATGTCGGCTTCTGGTATCCGTCGGTACTCGGTGATCAACGCC	2580
Db	2521	CCGAACATCCCGCAGATGTCGGCTTCTGGTATCCGTCGGTACTCGGTGATCAACGCC	2580
QY	2581	GCCAGCGTCTGACAGCTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
Db	2581	GCCAGCGTCTGACAGCTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
QY	2641	AACAAACACACATACATTAACAAACAACTCGGATCGAGGAGAGATTTCAGAAATTC	2700
Db	2641	AACAAACACACATACATTAACAAACAACTCGGATCGAGGAGAGATTTCAGAAATTC	2700
RESULT 10			
AR031992			
LOCUS	AR031992	7475 bp	DNA
DEFINITION	Sequence 1 from patent US 5866684.		
ACCESSION	AR031992		
VERSION	AR031992.1	GI:5946281	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 7475)		
AUTHORS	Atwood,M.Richard., Hurst,D.Nigel., Jones,P.Stephen.,		
TITLE	Key,P.Brittain., Raynham,T.Michael. and Wilson,F.Xavier.		
JOURNAL	Peptidyl inhibitors of viral proteases		
FEATURES	Patent: US 5866684-A 1 02-FEB-1999;		
	Location/Qualifiers		
source	1..7475		
	/organism="unknown"		
	/mol_type="unassigned DNA"		
ORIGIN			
Query Match	81.8%;	Score 2700;	DB 6; Length 7475;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 2700;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY	1	CCGACACCATCGAATGGTGTCAAAACCTTTTCGCGGTATGGCATGATGCGCCGGAAGAGA	60
Db	1	CCGACACCATCGAATGGTGTCAAAACCTTTTCGCGGTATGGCATGATGCGCCGGAAGAGA	60
QY	61	GTCAAATTCCAGGTGGTGAATGTCAAAACCAAGTATACAGTTATACAGTATGCGAGATATGCCG	120
Db	61	GTCAAATTCCAGGTGGTGAATGTCAAAACCAAGTATACAGTTATACAGTATGCGAGATATGCCG	120
QY	121	GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCACGTTTCTTGCAGAAA	180
Db	121	GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCACGTTTCTTGCAGAAA	180
QY	181	CGCGGCAAAAGTGAAGCGGCGATCGCGAGCTGAATTACATTCCCAACCGGTGGCAC	240

Qy	1321	TGTCGCTCAAGCGCGAC	CTCCGGTTCTCGATAATGTTTTTGGCCGACATCATAAACGGTT	1380
Db	1321	TGTCGCTCAAGCGCGAC	CTCCGGTTCTCGATAATGTTTTTGGCCGACATCATAAACGGTT	1380
Qy	1381	CTGGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCTGGCTCGTAAATGTTGGGA	1440	
Db	1381	CTGGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCTGGCTCGTAAATGTTGGGA	1440	
Qy	1441	ATTGTGAGCGGATAACAATTTTACACAGGAAACAGCCAGTCCGGTTTAGGTGTTTTTACAGA	1500	
Db	1441	ATTGTGAGCGGATAACAATTTTACACAGGAAACAGCCAGTCCGGTTTAGGTGTTTTTACAGA	1500	
Qy	1501	GCACTTCAACCAAGGACCATAGATTATGAAAACTGAAGAAGTAAACTGGTAATCTGG	1560	
Db	1501	GCACTTCAACCAAGGACCATAGATTATGAAAACTGAAGAAGTAAACTGGTAATCTGG	1560	
Qy	1561	ATTAAACGGCGATAAAGCTATAACCGTCTCGCTGAAGTCGGTAGAANAATTCGAGANAAGAT	1620	
Db	1561	ATTAAACGGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAGAANAATTCGAGANAAGAT	1620	
Qy	1621	ACCGGAATTAAGTACACCGTTGAGCATCCGGATAAACTGGAAGAGAAAATTCCCACAGGTT	1680	
Db	1621	ACCGGAATTAAGTACACCGTTGAGCATCCGGATAAACTGGAAGAGAAAATTCCCACAGGTT	1680	
Qy	1681	GCGGCAACTGGCGATGGCCCTGCACATATTCTTCTGGGCACACGACCGCTTTGGTGCGTAC	1740	
Db	1681	GCGGCAACTGGCGATGGCCCTGCACATATTCTTCTGGGCACACGACCGCTTTGGTGCGTAC	1740	
Qy	1741	GCTCAATCTGGCCCTGTTGGCTGAAATCAACCCCGGACAAAGGTTTCCAGGCAAGCTGTAT	1800	
Db	1741	GCTCAATCTGGCCCTGTTGGCTGAAATCAACCCCGGACAAAGGTTTCCAGGCAAGCTGTAT	1800	
Qy	1801	CCGTTTACCTGGGATGCGTAGCTTACAAACGGCAAGCTGATTGCTTACCCGATCGCTGTT	1860	
Db	1801	CCGTTTACCTGGGATGCGTAGCTTACAAACGGCAAGCTGATTGCTTACCCGATCGCTGTT	1860	
Qy	1861	GAAACGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAAACCCCGCAAAAACCTCGGAA	1920	
Db	1861	GAAACGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAAACCCCGCAAAAACCTCGGAA	1920	
Qy	1921	GAGATCCGGCGCTGGATAAAGACTGAAGCGAAAGGTAAAGCGCGCTGATGTTTCAAC	1980	
Db	1921	GAGATCCGGCGCTGGATAAAGACTGAAGCGAAAGGTAAAGCGCGCTGATGTTTCAAC	1980	
Qy	1981	CTGCAAGAACCGTACTTCACTGCGCGCTGATTCGCTGACGGGGGTATGCGTTCAAG	2040	
Db	1981	CTGCAAGAACCGTACTTCACTGCGCGCTGATTCGCTGACGGGGGTATGCGTTCAAG	2040	
Qy	2041	TATGAAAAACGGCAAGTAGTAGCATTTAAAGACGTTGGGCGTGGATAACGCTGGCGGAAAACGG	2100	
Db	2041	TATGAAAAACGGCAAGTAGTAGCATTTAAAGACGTTGGGCGTGGATAACGCTGGCGGAAAACGG	2100	
Qy	2101	GGTCTGACCTTCTGGTTGACCTGATTAAAAACAAACACATGAATGAGACACCGATTAC	2160	
Db	2101	GGTCTGACCTTCTGGTTGACCTGATTAAAAACAAACACATGAATGAGACACCGATTAC	2160	
Qy	2161	TCCATCGCAGAGCTGCGCTTTTAATAAGGCGGAAAACAGGATGACCATCAACGGCCCGTGG	2220	
Db	2161	TCCATCGCAGAGCTGCGCTTTTAATAAGGCGGAAAACAGGATGACCATCAACGGCCCGTGG	2220	
Qy	2221	GCATGGTCCAACTCGACACCAAGTGAATTTATGGTGTAAACGGTACTGCGCACCTTC	2280	
Db	2221	GCATGGTCCAACTCGACACCAAGTGAATTTATGGTGTAAACGGTACTGCGCACCTTC	2280	
Qy	2281	AAGGGTCAACCATCAAAACCGTTCTGTTGGCGTCTGAGCGCAGGTATTAACCGCGCCAGT	2340	
Db	2281	AAGGGTCAACCATCAAAACCGTTCTGTTGGCGTCTGAGCGCAGGTATTAACCGCGCCAGT	2340	
Qy	2341	CCGAACAAAGAGCTGGCNAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCGTG	2400	
Db	2341	CCGAACAAAGAGCTGGCNAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCGTG	2400	

Qy	2401	GAACGGGTTAATAAAGACAAACCCCTCGGTGCGGTAGCGCTGAAGCTTTACGAGGAAGAG	2460					
Db	2401	GAACGGGTTAATAAAGACAAACCCCTCGGTGCGGTAGCGCTGAAGCTTTACGAGGAAGAG	2460					
Qy	2461	TTGCGGAAAGATCCACGCTATTGCCGCCACCATGGAACGCCCAAGAAAGGTGAAATCATG	2520					
Db	2461	TTGCGGAAAGATCCACGCTATTGCCGCCACCATGGAACGCCCAAGAAAGGTGAAATCATG	2520					
Qy	2521	CCGAACATCCCGCAGATGTCCGCTTTCTTGGTATGTCGCTGCGTACTCGCGTGATCAACGCC	2580					
Db	2521	CCGAACATCCCGCAGATGTCCGCTTTCTTGGTATGTCGCTGCGTACTCGCGTGATCAACGCC	2580					
Qy	2581	GCCAGCGGTGCTCAGACTGTCGATGAAGCCCTGAAAGACGGCGAGACTAATTCGAGTCTG	2640					
Db	2581	GCCAGCGGTGCTCAGACTGTCGATGAAGCCCTGAAAGACGGCGAGACTAATTCGAGTCTG	2640					
Qy	2641	AACAAACAAACAATAACAATAACAACAACCTCGGATCGAGGGAAGATTTTCAGAATTC	2700					
Db	2641	AACAAACAAACAATAACAATAACAACAACCTCGGATCGAGGGAAGATTTTCAGAATTC	2700					
RESULT 11								
AR207294								
LOCUS	AR207294	7475 bp	DNA linear PAT 20-JUN-2002					
DEFINITION	Sequence 1 from patent US 6372883.							
ACCESSION	AR207294							
VERSION	AR207294.1	GI:21506162						
KEYWORDS								
SOURCE	Unknown.							
ORGANISM	Unknown.							
REFERENCE	Unclassified.							
AUTHORS	1 (bases 1 to 7475)							
TITLE	Atwood, M. Richard., Hurst, D. Nigel., Jones, P. Stephen.,							
	Kay, P. Brittain., Raynham, T. Michael. and Wilson, P. Xavier.							
	Antiviral medicaments							
	Patent: US 6372883-A 1 16-APR-2002;							
FEATURES	Location/Qualifiers							
	1..7475							
	/organism="unknown"							
ORIGIN	/mol_type="unassigned DNA"							
Query Match 81.8%; Score 2700; DB 6; Length 7475;								
Best Local Similarity 100.0%; Pred. No. 0;								
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;								
Qy	1	CCGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGAAGAGA	60					
Db	1	CCGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGAAGAGA	60					
Qy	61	GTCAATTACAGGTCGTGAACTGTGAACACAGTAAAGCTTATACGATGTCGAGAGTATCCG	120					
Db	61	GTCAATTACAGGTCGTGAACTGTGAACACAGTAAAGCTTATACGATGTCGAGAGTATCCG	120					
Qy	121	GTGTCTCTTATCAGACCGTTTCCCGCGTGTGTGAACACGAGCCAGCCACCGTTTCTCGAAAA	180					
Db	121	GTGTCTCTTATCAGACCGTTTCCCGCGTGTGTGAACACGAGCCAGCCACCGTTTCTCGAAAA	180					
Qy	181	CGCGGAAAAAGTGGAGCGCGGATGCGGAGCTGAAATTACATTCACCAACCGCGTGGCAC	240					
Db	181	CGCGGAAAAAGTGGAGCGCGGATGCGGAGCTGAAATTACATTCACCAACCGCGTGGCAC	240					
Qy	241	AACAACTGGCGGGCAACAGTCGTTGTTGATGTCGCTGTCACCTCCAGTCTGACCTGCG	300					
Db	241	AACAACTGGCGGGCAACAGTCGTTGTTGATGTCGCTGTCACCTCCAGTCTGACCTGCG	300					
Qy	301	ACGCGCGCTCGCAAAATTGTCCGCGCGATTAATCTCCGCGCGATCAACTGGGTGCCAGCG	360					
Db	301	ACGCGCGCTCGCAAAATTGTCCGCGCGATTAATCTCCGCGCGATCAACTGGGTGCCAGCG	360					
Qy	361	TGTTGGTTCGATGTTAGAACGAAGCGCGCTCGAAGCTGTAAAGCGCGGTGCACAATC	420					
Db	361	TGTTGGTTCGATGTTAGAACGAAGCGCGCTCGAAGCTGTAAAGCGCGGTGCACAATC	420					

QY 421 TTCTCGCGAAGCGCTCAGTGGCTGATCAATTAACTATCCGTGGATGACACGAGATGCCA 480
DB |||||
QY 421 TTCTCGCGAAGCGCTCAGTGGCTGATCAATTAACTATCCGTGGATGACACGAGATGCCA 480
DB |||||
QY 481 TTGCTGTGGAGCTCGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTGTACACAGA 540
DB |||||
QY 481 TTGCTGTGGAGCTCGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTGTACACAGA 540
DB |||||
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGCATCGGCGGTGGAGCATC 600
DB |||||
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGCATCGGCGGTGGAGCATC 600
DB |||||
QY 601 TGGTCGCATTTGGGTACACAGAAATTCGCGTGTAGCGGGCCCAATTAAGTTCTGTCTGG 660
DB |||||
QY 601 TGGTCGCATTTGGGTACACAGAAATTCGCGTGTAGCGGGCCCAATTAAGTTCTGTCTGG 660
DB |||||
QY 661 CGCGTCTCGCTCGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB |||||
QY 661 CGCGTCTCGCTCGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB |||||
QY 721 CGGAAACGGGAAGGACATGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
DB |||||
QY 721 CGGAAACGGGAAGGACATGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
DB |||||
QY 781 ATGAGGGCATCGTTCCCATCTCGCATGCTGGTGGCAACGATCGATGGCGTGGCGCGCAA 840
DB |||||
QY 781 ATGAGGGCATCGTTCCCATCTCGCATGCTGGTGGCAACGATCGATGGCGTGGCGCGCAA 840
DB |||||
QY 841 TGGCGCCATTACCGAGTCCGGGCTGGCGTGGTGGGATATCTCGTACTGGGATACG 900
DB |||||
QY 841 TGGCGCCATTACCGAGTCCGGGCTGGCGTGGTGGGATATCTCGTACTGGGATACG 900
DB |||||
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAACCATCAAAACAGGATTTTC 960
DB |||||
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAACCATCAAAACAGGATTTTC 960
DB |||||
QY 961 GCCTGTGGGCGAAACGAGCTGACCGCTTGTGTGCAACTCTTCAGGGCGAGCGGTGA 1020
DB |||||
QY 961 GCCTGTGGGCGAAACGAGCTGACCGCTTGTGTGCAACTCTTCAGGGCGAGCGGTGA 1020
DB |||||
QY 1021 AGGCAATCAGCTGTGTCGCGCTCTCACTGTGTAAGAAAGAAACCAACCTGCGCGCCCAATA 1080
DB |||||
QY 1021 AGGCAATCAGCTGTGTCGCGCTCTCACTGTGTAAGAAAGAAACCAACCTGCGCGCCCAATA 1080
DB |||||
QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCAATTAATGCAGCTGGCACGACAGGTTT 1140
DB |||||
QY 1141 CCGACTGGAAGCGGCGAGTGAGGCGNAACCAATTAATGTAGTTAGTCACTCATTTAG 1200
DB |||||
QY 1141 CCGACTGGAAGCGGCGAGTGAGGCGNAACCAATTAATGTAGTTAGTCACTCATTTAG 1200
DB |||||
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGCG 1260
DB |||||
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGCG 1260
DB |||||
QY 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGACAGTGTGTAATCACTGCATAATTCG 1320
DB |||||
QY 1321 TGTGCTCAAGGCGCATCTCCGTTCTGGATTAATGTTTTTGGCGCGACATCAATACCGGTT 1380
DB |||||
QY 1321 TGTGCTCAAGGCGCATCTCCGTTCTGGATTAATGTTTTTGGCGCGACATCAATACCGGTT 1380
DB |||||
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA 1440
DB |||||
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA 1440
DB |||||
QY 1441 ATTGTAGCGGATAACTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
DB |||||
QY 1441 ATTGTAGCGGATAACTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
DB |||||

QY 1501 GCATTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGTAAACTGTTAATCTGG 1560
DB |||||
QY 1501 GCATTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGTAAACTGTTAATCTGG 1560
DB |||||
QY 1561 ATTAACCGCGATAAAGGCTATACCGTCTCGTGAAGTCCGTAAGAAATTCGAGAAAGAT 1620
DB |||||
QY 1561 ATTAACCGCGATAAAGGCTATACCGTCTCGTGAAGTCCGTAAGAAATTCGAGAAAGAT 1620
DB |||||
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCCGATAAACTGGAAGAGAAATTCGCCACAGTT 1680
DB |||||
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCCGATAAACTGGAAGAGAAATTCGCCACAGTT 1680
DB |||||
QY 1681 GCGCAAACTGGCGATGCCCTGACATTTCTTTGGGCAACACGACCGCTTTTGGTGGCTAC 1740
DB |||||
QY 1741 GCTCAATCTGGCCTGTTGGCTGAATCACCCCGACAAAGGTTCCAGGACAACTGTAT 1800
DB |||||
QY 1741 GCTCAATCTGGCCTGTTGGCTGAATCACCCCGACAAAGGTTCCAGGACAACTGTAT 1800
DB |||||
QY 1801 CCGTTTACCTGGGATGCCGTACGTTACAAACGCAAGCTGATTGCTTACCCTGATCGCTGTT 1860
DB |||||
QY 1801 CCGTTTACCTGGGATGCCGTACGTTACAAACGCAAGCTGATTGCTTACCCTGATCGCTGTT 1860
DB |||||
QY 1861 GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCGCCCAAAACCTCGGAA 1920
DB |||||
QY 1861 GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCGCCCAAAACCTCGGAA 1920
DB |||||
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAACGCAAGGTAAGAGCGCGCTGATGTTCAAC 1980
DB |||||
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAACGCAAGGTAAGAGCGCGCTGATGTTCAAC 1980
DB |||||
QY 1981 CTGCAAGAACCGTACTTCACTGCGCTGATTGCTGACCGGGGTTATGCTTCAAG 2040
DB |||||
QY 1981 CTGCAAGAACCGTACTTCACTGCGCTGATTGCTGACCGGGGTTATGCTTCAAG 2040
DB |||||
QY 2041 TATGAAACCGCAAGTACGACATTAAGAGCGTGGCGTGGATAAACGCTGGCGGAAAGCG 2100
DB |||||
QY 2041 TATGAAACCGCAAGTACGACATTAAGAGCGTGGCGTGGATAAACGCTGGCGGAAAGCG 2100
DB |||||
QY 2101 GGTCTGACCTTCTGTTGACCTGATTAAGAAACAAACACATGATGACGACACCGGATAC 2160
DB |||||
QY 2101 GGTCTGACCTTCTGTTGACCTGATTAAGAAACAAACACATGATGACGACACCGGATAC 2160
DB |||||
QY 2161 TCCATCGCAGAAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACCGGCCCTGG 2220
DB |||||
QY 2161 TCCATCGCAGAAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACCGGCCCTGG 2220
DB |||||
QY 2221 GCATGTTCCAAACATCGACACAGCAAAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
DB |||||
QY 2221 GCATGTTCCAAACATCGACACAGCAAAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
DB |||||
QY 2281 AAGGTTCAACATCCAAACCGTTCTGTTGGCGTCTGAGCGCAGGTATTAACCGGCCAGT 2340
DB |||||
QY 2281 AAGGTTCAACATCCAAACCGTTCTGTTGGCGTCTGAGCGCAGGTATTAACCGGCCAGT 2340
DB |||||
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
DB |||||
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCG 2400
DB |||||
QY 2401 GAAGCGGTTAATAAGACAAACCGTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
DB |||||
QY 2401 GAAGCGGTTAATAAGACAAACCGTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
DB |||||
QY 2461 TTGCGGAAAGATCCAGTATTTGCCGCCACCATGGAACCCAGAAAGGTGAATCATG 2520
DB |||||
QY 2461 TTGCGGAAAGATCCAGTATTTGCCGCCACCATGGAACCCAGAAAGGTGAATCATG 2520
DB |||||
QY 2521 CCGAAACATCCCGCAGATGTCGCTTTCTGTTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
DB |||||
QY 2521 CCGAAACATCCCGCAGATGTCGCTTTCTGTTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
DB |||||
QY 2581 GCCAGCGGTGTCAGACTGTCTGATGAAGCCCTGAAAGACCGCGCAGACTAATTCGAGCTCG 2640
DB |||||

Db 2581 |||||GCCACGGTCTGCAGCTGTCGATGAAGCCCTGAAAGAGCGCGACACTAAATTCGAGCTCG 2640

Qy 2641 AACAAACAACAACAATAACAATAACAACCTCGGGATCGAGGAGGATTTTCAGAAATTC 2700

Db 2641 AACAAACAACAACAATAACAATAACAACCTCGGGATCGAGGAGGATTTTCAGAAATTC 2700

RESULT 12

AX377532

LOCUS AX377532 9191 bp DNA linear PAT 18-MAR-2002

DEFINITION Sequence 9 from Patent WO0212553.

ACCESSION AX377532

VERSION AX377532.1 GI:19573718

KEYWORDS synthetic construct

SOURCE synthetic construct

ORGANISM other sequences; artificial sequences.

REFERENCE 1

AUTHORS Kappel,A., Polakowski,T., Pignot,M., Windhab,N., Behrensdoerf,H. and Muth,J.

TITLE Method for detecting mutations in nucleotide sequences

JOURNAL Patent: WO 0212553-A 9 14-FEB-2002;

FEATURES

source Location/Qualifiers

1..9191

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

/note="pMALc2x-mutS Plasmid"

ORIGIN

Query Match 81.8%; Score 2698; DB 6; Length 9191;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2701; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CCGACACCATCAATGGTGCAGAAACCTTTCGGGGTATGGCATGTAGCGCCCGGAAGAGA 60

Db 1 CCGACACCATCAATGGTGCAGAAACCTTTCGGGGTATGGCATGTAGCGCCCGGAAGAGA 60

Qy 61 GTCAATTACAGGTGGTGAATGTGAACACAGTAACTTATACCATGTGCGAGAGTATGCCG 120

Db 61 GTCAATTACAGGTGGTGAATGTGAACACAGTAACTTATACCATGTGCGAGAGTATGCCG 120

Qy 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACAGTAACTTATACCATGTGCGAGAGTATGCCG 180

Db 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACAGTAACTTATACCATGTGCGAGAGTATGCCG 180

Qy 181 CCGGGGAAAAAGTGAAGCGCGATGGCGGAGCTGAAATTCACATTCGCCAACCCGCTGGCAC 240

Db 181 CCGGGGAAAAAGTGAAGCGCGATGGCGGAGCTGAAATTCACATTCGCCAACCCGCTGGCAC 240

Qy 241 AACAACTGCGGGCAACAGTCTGCTCATTTGGCGTTGCGACCTCCAGTCTGCGCCCTGC 300

Db 241 AACAACTGCGGGCAACAGTCTGCTCATTTGGCGTTGCGACCTCCAGTCTGCGCCCTGC 300

Qy 301 ACGCGCGTTCGCAAAATTTGCGGGGATTAATCTCGCCCGATCAACTGGGTGCGCAGCG 360

Db 301 ACGCGCGTTCGCAAAATTTGCGGGGATTAATCTCGCCCGATCAACTGGGTGCGCAGCG 360

Qy 361 TGGTGTGTCGATGTGATAGAACGAGCGGCTGAAAGCTGTAAAGCGGCGTGCACAATC 420

Db 361 TGGTGTGTCGATGTGATAGAACGAGCGGCTGAAAGCTGTAAAGCGGCGTGCACAATC 420

Qy 421 TTCTCGCGCAACCGCTCAGTGGGTGATCAATTAATCTCCGCTGGATGACAGGATGCCA 480

Db 421 TTCTCGCGCAACCGCTCAGTGGGTGATCAATTAATCTCCGCTGGATGACAGGATGCCA 480

Qy 481 TTGCTGTGGAAGCTGCTGCATTAATGTTCGGCGTTATTTCTTGATGTCTCTGACCGA 540

Db 481 TTGCTGTGGAAGCTGCTGCATTAATGTTCGGCGTTATTTCTTGATGTCTCTGACCGA 540

Qy 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600

Db 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600

Qy 601 TGGTGCATTTGGGTCCACAGCAAAATCGCGCTGTAGTGGGGCCCAATTAAGTTCTGTCTCGG 660

Db 601 TGGTGCATTTGGGTCCACAGCAAAATCGCGCTGTAGTGGGGCCCAATTAAGTTCTGTCTCGG 660

Qy 661 CGCGTCTCGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGCATAG 720

Db 661 CGCGTCTCGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGCATAG 720

Qy 721 CGGAACGGGAAGGACCTGGAGTGCATGTCCGCTTAAACAAACCAATTCAGAAATTCGCTGA 780

Db 721 CGGAACGGGAAGGACCTGGAGTGCATGTCCGCTTAAACAAACCAATTCAGAAATTCGCTGA 780

Qy 781 ATGAGGGCATGTTCCCACTCGCATGCTGGTTGCCAACGATCAGATGCGCTGGCGCGCAA 840

Db 781 ATGAGGGCATGTTCCCACTCGCATGCTGGTTGCCAACGATCAGATGCGCTGGCGCGCAA 840

Qy 841 TGGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGGATATCTCGGTAGTGGGATAG 900

Db 841 TGGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGGATATCTCGGTAGTGGGATAG 900

Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960

Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960

Qy 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCGCCAGGCGTGA 1020

Db 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCGCCAGGCGTGA 1020

Qy 1021 AGGGCAATCAGCTGTTGCCCTCTCATCTGGTGAAGAAAAACACCTTGGCGCCCAATA 1080

Db 1021 AGGGCAATCAGCTGTTGCCCTCTCATCTGGTGAAGAAAAACACCTTGGCGCCCAATA 1080

Qy 1081 CGCAAAACCGCTCTCCCGCGTGGCGGATTCATTAAATGAGTGGCAGCAGCAGGTTT 1140

Db 1081 CGCAAAACCGCTCTCCCGCGTGGCGGATTCATTAAATGAGTGGCAGCAGCAGGTTT 1140

Qy 1141 CCGACTCGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTACTCATTTAG 1200

Db 1141 CCGACTCGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTACTCATTTAG 1200

Qy 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGGTCGACCAATGCTTCTGGCG 1260

Db 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGGTCGACCAATGCTTCTGGCG 1260

Qy 1261 TCAGGACCCATCGGAAGCTGTGTATGGCTGTGCGAGTCTGTAATCATCTGCATTAATTCG 1320

Db 1261 TCAGGACCCATCGGAAGCTGTGTATGGCTGTGCGAGTCTGTAATCATCTGCATTAATTCG 1320

Qy 1321 TGTGCTCAAGCGCACCTCCCGTCTCGGATTAATGTTTTTGGCGCGACATCAATAACGGTT 1380

Db 1321 TGTGCTCAAGCGCACCTCCCGTCTCGGATTAATGTTTTTGGCGCGACATCAATAACGGTT 1380

Qy 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440

Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440

Qy 1441 ATTGTAGCGGATTAACAATTTTCAACAGGAACACCGCAGTCCGTTTAGTGTTCACGA 1500

Db 1441 ATTGTAGCGGATTAACAATTTTCAACAGGAACACCGCAGTCCGTTTAGTGTTCACGA 1500

Qy 1501 GCACCTTCAACCAAGGACCATAGATTAATGAAAACTGAAGAGGTAAACTGGTAATCTCG 1560

Db 1501 GCACCTTCAACCAAGGACCATAGATTAATGAAAACTGAAGAGGTAAACTGGTAATCTCG 1560

Qy 1561 ATTAACGGCGATAAAGGCTATAACCGCTCGCTGAAGTCCGTAAGAAATTCGAGAAAGAT 1620

Db 1561 ATTAACGGCGATAAAGGCTATAACCGCTCGCTGAAGTCCGTAAGAAATTCGAGAAAGAT 1620

Qy 1621 ACCCGAATTAAGTCAACCGTTGAGCATCCGGATAACTGGGAAGAGAAATTCACAGGTT 1680

Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAAATCCACAGGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GGTCAATCTGGCTGTGGCTGAATATCAACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GGTCAATCTGGCTGTGGCTGAATATCAACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Qy 1801 CGTTTACTCGGATGGCGTACGTTTACACCGCAAGCTGATGCTTACCGATCGCTGTT 1860
Db 1801 CGTTTACTCGGATGGCGTACGTTTACACCGCAAGCTGATGCTTACCGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCGCTGGATAAGAACTGAAGCGAAGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAGAACTGAAGCGAAGTAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTCTCACCTGGCGCTGATGCTGCTGACGGGGTTATCGCTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTCTCACCTGGCGCTGATGCTGCTGACGGGGTTATGCGGTTCAAG 2040
Qy 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGCGAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGCGAAGCG 2100
Qy 2101 GGTCTGACCTTCTGGTTGACCTGATTTAAAAACAAACATCATGATGACAGACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTTAAAAACAAACATCATGATGACAGACCGGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTAATAAGCGCAACACAGCGATGACCATCAACGGCGCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAGCGCAACACAGCGATGACCATCAACGGCGCGCTGG 2220
Qy 2221 GCATGGTCCAAACATCGACACACGCAAAAGTGAATTTATGGTGTAAACGGTACTCGGACCTTC 2280
Db 2221 GCATGGTCCAAACATCGACACACGCAAAAGTGAATTTATGGTGTAAACGGTACTCGGACCTTC 2280
Qy 2281 AAGGTCACCATCCAAACCGTTTGGCGTGTGACGCGCAGGTATTAACCGCGCCAGT 2340
Db 2281 AAGGTCACCATCCAAACCGTTTGGCGTGTGACGCGCAGGTATTAACCGCGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACATCTGCTGACTGATGAGGTCG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACATCTGCTGACTGATGAGGTCG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGCCGTAGCGCTGAAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGCCGTAGCGCTGAAAGTCTTACGAGGAAG 2460
Qy 2461 TTGCGGAAGATCCACGTTATGCGCCACCATGAAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGCGGAAGATCCACGTTATGCGCCACCATGAAACCGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATCCGCTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATCCGCTGCGTACTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGGTGTCAGCTGTCGATGAAGCCCTGAAACCGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGTCGAGCTGTCGATGAAGCCCTGAAACCGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGATTTCAGATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGATTTCAGATTC 2700
Qy 2701 GGATCC 2706
Db 2701 GGATCC 2706

RESULT 13
AF097412
LOCUS 6724 bp DNA circular SYN 28-JAN-2002
DEFINITION Expression vector pMBP-parallel1, complete sequence.
ACCESSION AF097412
VERSION AF097412.1 GI:3983120
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 6724)
AUTHORS Sheffield, P., Garrard, S. and Derewenda, Z.
TITLE Overcoming expression and purification problems of RhoGDI using a family of 'parallel' expression vectors
JOURNAL Protein Expr. Purif. 15 (1), 34-39 (1999)
MEDLINE 99150479
PUBMED 10024467
REFERENCE
2 (bases 1 to 6724)
AUTHORS Sheffield, P.J., Garrard, S.M. and Derewenda, Z.S.
TITLE Direct Submission
JOURNAL Submitted (05-OCT-1998) Molecular Physiology and Biological Physics, University of Virginia, 4215 Jordan Hall, 1300 Jefferson Park Avenue, Charlottesville, Virginia 22908, USA
FEATURES
Location/Qualifiers
1..6724
/organism="Expression vector pMBP-parallel1"
/mol_type="other DNA"
/db_xref="taxon:84479"
204..1163
/gene="lacIq"
204..1163
/gene="lacIq"
/codon_start=1
/transl_table=11
/product="lac repressor"
/protein_id="AAC83812.1"
/db_xref="GI:3983121"
/translations="MAELNYIPNRVAQQLAGKQSLIGVATSLALHAPSOIVAAIKS
RADQLGASVVSVMSVERSGVEACKAAVHNLLAQRVSLIINYPDDQDAIAVEACTNV
PALFLDVSQTPINSIIFSHEDGTRGLVHVALGHQQLIALLAGPLSSVARSRLAGW
HKYLTNRQIPIAEREGDWSAMSGFOOTQMLNEGIQVPTAMLVANDQMALGRAMRAITE
SGLRVGADISVVGDDTSDSSCYIPPLTTIKODFRLLGOTSDVRLILQLSQGQAVKGNQ
LLPVSILVKRKTTLAPNTQTASPRALADSLMQLARQVSRLESG"
1453..1481
/note="tac; inducible promoter"
1528..2994
/genes="male"
1528..2994
/genes="male"
/codon_start=1
/transl_table=11
/product="maltose binding protein"
/protein_id="AAC83813.1"
/db_xref="GI:3983122"
/translations="MKIERGLVITWINGDKYNGLAEVGKPEKDTGIVTVHPDKL
EKFPVAATGDPDIIFWAHDPRPGYAGSGLLAETPKAFQDKLYLPFTWDVRYNG
KLIAYIAVEALSLIYNKOLLPPPKTWEEIPALDKELKAKGSALFNLQEPYFTWP
LIAADGGYAPKVENGYKIDKVDVNDAGAKAGTLFLVDLTKNKNMADTDYSABEAF
NKGETAMTTPGPAWSNIDTSKYNVGTIVLTPFKGQSPKFPVGVLSAGINAASPNKEL
AKFELNYLLTDEGLEAVNKKPLGAVALKSYEEELAKDPRIATAATMENAKGIMPN
POMSAFWAVRTAVINAAGROTVDEALKDAQTNSNNNNNNNNNNNNLIGDITENLYP
QAMDPKPLRRRAQLVRPLNLEPAVSRHAPVPSIALAVLVLRQRDWENPGVTQLNRL
AAHPSPASWENSEEARTDRPSQQLRSNLNGEWQLGCFGG"
2633..2688
/genes="male"
/note="protein spacer region"
2689..2709
/genes="male"
/note="Tobacco etch virus protease cleavage site"
2711..2810
/genes="male"

gene	/note="multiple cloning site"			
promoter	3426.. .4356			
	/gene="bla"			
	3426.. .3454			
	/gene="bla"			
	3496.. .4356			
CDS	/gene="bla"			
	/codon_start=1			
	/transl_table=11			
	/product="beta-lactamase"			
	/protein_id="AAC83814.1"			
	/db_xref="GI:3983123"			
	/translation="MSIQHFRVALIIPFAAFCLPVFAHPETLVKVKDAEDQLGARVGY			
	IEDLNSKILDESFPBERSPMSTFKVLLCGAVLSVRDAGEQGLRRIHYSONDLVE			
	YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLTIIGPKELTAFIHNMGDHDVTRL			
	DRWPELNEAIPNDRSDTTPVMAATILRLKLTGELLTLASRQQLIDNWEADKVGAPL			
ORIGIN	LRSLAPAGWFIADKSGAGBSRGI I AALGPDGKPSRIWVIYTTGSGQATMDERNQIIA			
	EIGASLIKHW"			
Query Match	81.1%;	Score 2676.2;	DB 12;	Length 6724;
Best Local Similarity	99.9%;	Pred. No. 0;		
Matches 2678;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;
QY	1	CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCCGCCGGAAGAGA	60	
Db	1	CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCCGCCGGAAGAGA	60	
QY	61	GTCAATTACAGGTGGTGAATGTGAACACAGTAACTTATACGATGTCGCAGAGTATGCCG	120	
Db	61	GTCAATTACAGGTGGTGAATGTGAACACAGTAACTTATACGATGTCGCAGAGTATGCCG	120	
QY	121	GTGTCTCTTATCAGACCGGTTTCCCGCGTGGTGAACCGAGCCAGCCACGTTTCTGCGAAA	180	
Db	121	GTGTCTCTTATCAGACCGGTTTCCCGCGTGGTGAACCGAGCCAGCCACGTTTCTGCGAAA	180	
QY	181	CGCGGAAAAAGTGGAAAGCGGCGATGGCGGAGCTGAAATTACATTCGCAACCCGCTGGC	240	
Db	181	CGCGGAAAAAGTGGAAAGCGGCGATGGCGGAGCTGAAATTACATTCGCAACCCGCTGGC	240	
QY	241	ACAACTCGCGGCAACAGCTGTGCTGATTTGGGTGGCCACCTCCAGCTGGCCCTGC	300	
Db	241	ACAACTCGCGGCAACAGCTGTGCTGATTTGGGTGGCCACCTCCAGCTGGCCCTGC	300	
QY	301	ACGCGCGTTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG	360	
Db	301	ACGCGCGTTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG	360	
QY	361	TGGTGGTGTGATGTAGAAAGAGCGGCTGCAAGCCTGTAAAGCGGCGGTGCACAATC	420	
Db	361	TGGTGGTGTGATGTAGAAAGAGCGGCTGCAAGCCTGTAAAGCGGCGGTGCACAATC	420	
QY	421	TTCTCGCGCAACGCTGAGTGGGCTGATCACTTAATCTCCGCTGGATGACAGGATGCCA	480	
Db	421	TTCTCGCGCAACGCTGAGTGGGCTGATCACTTAATCTCCGCTGGATGACAGGATGCCA	480	
QY	481	TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA	540	
Db	481	TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA	540	
QY	541	CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600	
Db	541	CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600	
QY	601	TGGTCCGATTTGGGTACCCAGCAATTCGGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG	660	
Db	601	TGGTCCGATTTGGGTACCCAGCAATTCGGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG	660	
QY	661	CGCGTCTGGCTGTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720	
Db	661	CGCGTCTGGCTGTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720	
QY	721	CGGAACGGGAAGCGGACTGGAGTGCATGTCCGGTTTTTCACAAACCATGCAATGCTGA	780	

721	CGGAACGGGAAGCGGACTGGAGTGCATGTCCCGTTTTTCACAAACCATGCAAAATGCTGA	780
781	ATGAGGGCATCGTTTCCCACTCGCATGCTGTGTGCCAACGATCAGATGGCGCTGGCGCAA	840
791	ATGAGGGCATCGTTTCCCACTCGCATGCTGTGTGCCAACGATCAGATGGCGCTGGCGCAA	840
841	TGCGCGCATTTACCGAGTCCGGGCTGGCGTTGGTGGCGATATCTCGGTAGTGGGATACG	900
841	TGCGCGCATTTACCGAGTCCGGGCTGGCGTTGGTGGCGATATCTCGGTAGTGGGATACG	900
901	ACGATACCGGAAGACAGCTCATGTTTATATCCCGCGTTTAAACACCATCAAAACAGATTTTC	960
901	ACGATACCGGAAGACAGCTCATGTTTATATCCCGCGTTTAAACACCATCAAAACAGATTTTC	960
961	GCCTGCTGGGGCAAAACAGCGGTGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
961	GCCTGCTGGGGCAAAACAGCGGTGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
1021	AGGGCAATCAGCTGTTGCCGCTCTCACTGTGTAAGAAAGAAACACCGCTGGCGCCCAATA	1080
1021	AGGGCAATCAGCTGTTGCCGCTCTCACTGTGTAAGAAAGAAACACCGCTGGCGCCCAATA	1080
1081	CGCAAAACCGCTCTCCCGCGGCTTGCGCATTCATTAATGACAGCTGGCACGACGAGTTT	1140
1081	CGCAAAACCGCTCTCCCGCGGCTTGCGCATTCATTAATGACAGCTGGCACGACGAGTTT	1140
1141	CCGACTTGGAAAGCGGCGAGTGAGCGCAACCGCAATTAATGTGAGTTAGTCTCACTATTAG	1200
1141	CCGACTTGGAAAGCGGCGAGTGAGCGCAACCGCAATTAATGTGAGTTAGTCTCACTATTAG	1200
1201	GCACAATTTCTCATGTTTGAACAGCTTATCATCGACTGCAACGATGCTTCTGGCG	1260
1201	GCACAATTTCTCATGTTTGAACAGCTTATCATCGACTGCAACGATGCTTCTGGCG	1260
1261	TCAGGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAAGTGTAAATCACTGCATATTCTG	1320
1261	TCAGGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAAGTGTAAATCACTGCATATTCTG	1320
1321	TGTGCTCAAGCGGCACCTCCCGTTCTGGATAATATGTTTTTTCGCGCCGACATCATAAACG	1380
1321	TGTGCTCAAGCGGCACCTCCCGTTCTGGATAATATGTTTTTTCGCGCCGACATCATAAACG	1380
1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA	1440
1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA	1440
1441	ATTGTAGCGGATTAACAAATTTACAGAGGAAACAGCCAGTCCGTTTAGGTTGTTTTTTCAC	1500
1441	ATTGTAGCGGATTAACAAATTTACAGAGGAAACAGCCAGTCCGTTTAGGTTGTTTTTTCAC	1500
1501	GCACTTTACCAACGAAGGACCATAGATTATGAAGAACTGAAGAGGTAAATCGGTAATCTGG	1560
1501	GCACTTTACCAACGAAGGACCATAGATTATGAAGAACTGAAGAGGTAAATCGGTAATCTGG	1560
1561	ATTTAACCGCGATAAAGGCTATAACCGGCTCGCTGAACTCGGTGAAGAAATTCGAGAAGAT	1620
1561	ATTTAACCGCGATAAAGGCTATAACCGGCTCGCTGAACTCGGTGAAGAAATTCGAGAAGAT	1620
1621	ACCGGAATTAAGTTCACCGTTGAGCATCCGGATAACTGGAAGAGAAATTCGAGAAGAT	1680
1621	ACCGGAATTAAGTTCACCGTTGAGCATCCGGATAACTGGAAGAGAAATTCGAGAAGAT	1680
1681	GGGGCAACTGGCGATGGCCCTGACATTTCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
1681	GGGGCAACTGGCGATGGCCCTGACATTTCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
1741	GCTCAATCTGGCCTGTTGGCTGAAATCAACCCCGGACAAAGCGTTTCCAGGCAAGCTGTAT	1800
1741	GCTCAATCTGGCCTGTTGGCTGAAATCAACCCCGGACAAAGCGTTTCCAGGCAAGCTGTAT	1800
1801	CCGTTTACCTGGGATGGCGTACGTTTCAACGGCAAGCTGATTTGCTTACCCTGCTGTTT	1860

Db 1801 CCGTTTAACTGGGATGCGGTACGTTTACACGGCAAGCTGATTGCTTACCCGATCGTGT 1860
QY 1861 GAAGCGTTATCGTGTATTTATACAAAGATCTGCTCGCAACCCGCCAAAACCTGGGA 1920
Db 1861 GAAGCGTTATCGTGTATTTATACAAAGATCTGCTCGCAACCCGCCAAAACCTGGGA 1920
QY 1921 GAGATCCCGCGCTCGATGAAGAACTGAAGCGAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTCGATGAAGAACTGAAGCGAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCTGCTTACCTGGCGCTGATGTTGCTGACGGGGTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACCTGCTTACCTGGCGCTGATGTTGCTGACGGGGTTATCGGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGAAGCGTGGCGTGGATAAGCGTGGCGGAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGAAGCGTGGCGTGGATAAGCGTGGCGGAAGCG 2100
QY 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAAAAACAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAAAAACAACACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTAATAAGCGCAAGCAAGCGATGACCATCAAGCGCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAGCGCAAGCAAGCGATGACCATCAAGCGCCGTGG 2220
QY 2221 GCATGGTCCAACTCGACACACAGCAAGTGAATATGTTGTAACGGTACTGCCACCTTC 2280
Db 2221 GCATGGTCCAACTCGACACACAGCAAGTGAATATGTTGTAACGGTACTGCCACCTTC 2280
QY 2281 AAGGTCAACCATCAAAACCGTTCTGTTGGCTGTGAGCGAGGTATTAAACGCCCAAGT 2340
Db 2281 AAGGTCAACCATCAAAACCGTTCTGTTGGCTGTGAGCGAGGTATTAAACGCCCAAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAAGCGTTAATAAGACAAACCGCTGGGTGCCGTAGCGCTGAAAGTCTTACGAGGAAG 2460
Db 2401 GAAAGCGTTAATAAGACAAACCGCTGGGTGCCGTAGCGCTGAAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAGATCCACGTATTTGCCGCCACATGGAAGACCCCAAGAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGTATTTGCCGCCACATGGAAGACCCCAAGAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGATGATGCGTGTACTGCGGTGATCAAGGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGATGATGCGTGTACTGCGGTGATCAAGGCC 2580
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACACAAACAATAAATAAACAACCAACCTCGGGATCGA 2681
Db 2641 AACACAAACAATAAATAAACAACCAACCTCGGGATCGA 2681

RESULT 14
AX378208
LOCUS AX378208 6806 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 194 from Patent WO0206306.
ACCESSION AX378208
VERSION AX378208.1 GI:19574061
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
AUTHORS Yan, R., Tomasselli, A.G., Gurney, M.E., Emmons, T.L., Bienkowski, M.J.
and Heinrikson, R.L.
TITLE Substrates and assays for g(b)-secretase activity

JOURNAL Patent: WO 0206306-A 194 24-JAN-2002;
PHARMACIA & UPJOHN COMPANY (US)
FEATURES location/Qualifiers
source 1. 6806
/organism="unidentified"
/mol type="unassigned DNA"
/db xref="taxon:32644"
/note="fusion protein comprising a maltose binding protein
with 125 amino"

ORIGIN

Query Match 79.9%; Score 2636.8; DB 6; Length 6806;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2638; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCGACACATCGAATGTGCAAAACCTTTCCGCGTATGGCATGATAGCCCGGAAGAGA 60
Db 1 CCGACACATCGAATGTGCAAAACCTTTCCGCGTATGGCATGATAGCCCGGAAGAGA 60
QY 61 GTCAATTCAGGTTGCAATGTGAAACCACTTAACGTTATACGATGTCGAGAGTATGCCG 120
Db 61 GTCAATTCAGGTTGCAATGTGAAACCACTTAACGTTATACGATGTCGAGAGTATGCCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACCAAGCGCCAGCTTTTTCGCGAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACCAAGCGCCAGCTTTTTCGCGAAA 180
QY 181 CGCGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTAATTCACATCCCGGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTAATTCACATCCCGGTGGCAC 240
QY 241 AACAACTGGCGGCAAAACAGTCTGCTGATTTCCGCGTTCGACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAAAACAGTCTGCTGATTTCCGCGTTCGACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCCGCTCGCAAAATTTGTCGCGCGATTAATAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCCGCTCGCAAAATTTGTCGCGCGATTAATAATCTCGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGTGTGTGTCGATGTTAGAACGAGCGGCTCGAAGCTGTAAAGCGCGGTGGCAATC 420
Db 361 TGTGTGTGTCGATGTTAGAACGAGCGGCTCGAAGCTGTAAAGCGCGGTGGCAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTCGCGTGGATGACCAAGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTCGCGTGGATGACCAAGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGACATAATTTTCCGCGTTATTTCTTGTATGTTCTTGACCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGACATAATTTTCCGCGTTATTTCTTGTATGTTCTTGACCAGA 540
QY 541 CACCCATCAACAGTATTATTTCTCCCATGAGACGGTACGGACTGGGGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAGACGGTACGGACTGGGGGTGGAGCATC 600
QY 601 TGTGTCATTGGGTACACGCAAAATCGCGTGTAGCGGCGCCATTAAGTTCTGTCTCGG 660
Db 601 TGTGTCATTGGGTACACGCAAAATCGCGTGTAGCGGCGCCATTAAGTTCTGTCTCGG 660
QY 661 CGGCTCTGGCTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGGCTCTGGCTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGCGGACTGGAGTCCATGTCGCGTTTTCACAAACCAACATGCAATCTGA 780
Db 721 CGGAACGGGAAGCGGACTGGAGTCCATGTCGCGTTTTCACAAACCAACATGCAATCTGA 780
QY 781 ATGAGGCAATCGTTCCCATGCGATGCTGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
Db 781 ATGAGGCAATCGTTCCCATGCGATGCTGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
QY 841 TGGCGCCATTACCGAGTCCGGCTGGCGTGGTGGGATATCTCGTAGTGGGATACG 900
Db 841 TGGCGCCATTACCGAGTCCGGCTGGCGTGGTGGGATATCTCGTAGTGGGATACG 900

Db 841 TGGCGCCATTACCGAGTCCGGCTCGCGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAACATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAACATCAAAACAGGATTTTC 960
Qy 961 GCCTGCTGGGGCAACACGAGCGGACCGCTTCTGCACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGGCAACACGAGCGGACCGCTTCTGCACTCTCTCAGGGCCAGGCGGTGA 1020
Qy 1021 AGGGAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAACCAACCTTGGCGCCCAATA 1080
Db 1021 AGGGAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAACCAACCTTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGGACAGCAGGTTT 1140
Db 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGGACAGCAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG 1200
Qy 1201 GCACAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCAACCAATGCTTCTGGCG 1260
Db 1201 GCACAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCAACCAATGCTTCTGGCG 1260
Qy 1261 TCAGCAGCAATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAATCACTGATCAATTCG 1320
Db 1261 TCAGCAGCAATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAATCACTGATCAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGCGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGCGACATCATACGGTT 1380
Qy 1381 CTGGCAATATCTGAATGAGCTGTGACAATTAATCATCGGCTCGTATAATGTGGA 1440
Db 1381 CTGGCAATATCTGAATGAGCTGTGACAATTAATCATCGGCTCGTATAATGTGGA 1440
Qy 1441 ATTGTAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCGGTTAGGTGTTTTTCAGGA 1500
Db 1441 ATTGTAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCGGTTAGGTGTTTTTCAGGA 1500
Qy 1501 GCACCTTCAACACAGGACCATAGATTATGAACCTGAAGAGGTAAACTGATATCTGG 1560
Db 1501 GCACCTTCAACACAGGACCATAGATTATGAACCTGAAGAGGTAAACTGATATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACGCTCTGCTGAAGTCGTAAGAAATTCAGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGCTCTGCTGAAGTCGTAAGAAATTCAGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATTCGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATTCGGAAGAGAAATTCACAGGTT 1680
Qy 1681 GGGCAATCGGCGATGGCCCTGACATATCTTCTGGGACACAGCCGCTTGGTGCGCTAC 1740
Db 1681 GGGCAATCGGCGATGGCCCTGACATATCTTCTGGGACACAGCCGCTTGGTGCGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGCTGAATCAACCGGACAAAGCGTTCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAATCAACCGGACAAAGCGTTCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCCGCTACGTTTACACGCGCAAGCTGATGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCCGCTACGTTTACACGCGCAAGCTGATGCTTACCCGATCGCTGT 1860
Qy 1861 GAAGGCTTATCGCTGATTTTAAACAAAGATCTGCTGGCGAAACCGCCCAAAACCTGGGAA 1920
Db 1861 GAAGGCTTATCGCTGATTTTAAACAAAGATCTGCTGGCGAAACCGCCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTCGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTCGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980

Qy 1981 CTGCAAGAACCGTACTTACCTTGGCCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTTGGCCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAACGCGCAAGTACGACATTAAGACGCTGGGCGTGGATAAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGCGCAAGTACGACATTAAGACGCTGGGCGTGGATAAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAAAACAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAAAACAACACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGAAGAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACGCGCCGTTG 2220
Db 2161 TCCATCGAAGAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACGCGCCGTTG 2220
Qy 2221 GCATGCTCAACATCGACACGCAAGCAAGTAAATGCTGTAACCGGTACTGCCGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACGCAAGCAAGTAAATGCTGTAACCGGTACTGCCGACCTTC 2280
Qy 2281 AAGGTCACACCATCAAAACCGTTTGGTGGCTGCTGACGCGAGGTATTAAACGCCCCAGT 2340
Db 2281 AAGGTCACACCATCAAAACCGTTTGGTGGCTGCTGACGCGAGGTATTAAACGCCCCAGT 2340
Qy 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGCTAGCGCTGAGCTGAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGCTAGCGCTGAGTCTTACGAGGAAG 2460
Qy 2461 TTGCGAAAGATCCACGATTATGCCGCCACCACTGAAAAACGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGCGAAAGATCCACGATTATGCCGCCACCACTGAAAAACGCCAGAAAGGTGAAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGCTCGCTTTTGGTATGCCGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCCGCAGATGCTCGCTTTTGGTATGCCGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Qy 2581 GCCAGCGTCTGACACTGCTGATGAGCTGTAAGAGCGCGAGACTTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACACTGCTGATGAGCTGTAAGAGCGCGAGACTTAATTCGAGCTCG 2640

RESULT 15
AF031088 6706 bp DNA circular SYN 07-MAY-2001
LOCUS Shuttle vector pMAL-pIII, complete sequence.
DEFINITION AF031088
ACCESSION AF031088
VERSION AF031088.1 GI:2623821
KEYWORDS Shuttle vector pMAL-pIII
SOURCE Shuttle vector pMAL-pIII
ORGANISM Shuttle vector pMAL-pIII
other sequences; artificial sequences; vectors.
REFERENCE 1 (bases 1 to 6706)
AUTHORS Zwick,M.B., Bonnycastle, L.L., Noren, K.A., Venturini, S., Leong, E.,
Barbas, C.F. III, Noren, C.J. and Scott, J.K.
TITLE The maltose-binding protein as a scaffold for monovalent display of
peptides derived from phage libraries
JOURNAL Anal. Biochem. 264 (1), 87-97 (1998)
MEDLINE 99002881
PUBMED 9784192
REFERENCE 2 (bases 1 to 6706)
AUTHORS Noren, K.A. and Noren, C.J.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-1997) New England Biolabs, 32 Tozer Road,
Beverly, MA 01915, USA
FEATURES
source
1..6706
/organism="Shuttle vector pMAL-pIII"
/mol_type="genomic DNA"

```

/db_xref="taxon:69327"
/Note="derived from commercially available vector pMal-p2
(New England Biolabs)"
15..1163
/gene="laci"
/Note="E. coli gene"
15..46
/gene="laci"
/Note="transcriptional repressor"
/codon_start=1
/transl_table=11
/product="lac repressor"
/protein_id="AAB86558.1"
/db_xref="GI:2623822"
/translation="MKPVTLYDAEYAGVSYQTVSRVNVQASHVSAKREKVEAMAB
LNYIPNRVAQQLAGKQSLIGVATSSLAHAPSQIVAAIKSRADOLGASVSVSVRS
GVKCAAVHNLQRVSGLLIINPLDDQDAIAVEAACTNVPALFLVSDQTPINSII
FSHEDGTRLGHVLAHQHQAIIALLAGLSVSARLRLAGWHKYLTRNQIPIAEREG
DWSAMSGPQTMQMLNEGIVPTAMLVANDQMALGMRAITBSGLRVGADI.SVVGDDT
EDSCYIPPLTTIKQDFRLIGQTSVDRLLQLSQCAVKGNQLLPVSLVKRKTLLAPNT
QTASPRALADSLMQLARQVSRLESQ"
1406..1433
/promoter
1528..2976
/gene="malE-lacZa fusion"
/Note="B. coli gene fusion"
1528..2976
/gene="malB-lacZa fusion"
/Note="MBP-lacZa; MBP modified to include phage M13
secretion signal sequence"
/codon_start=1
/transl_table=11
/product="maltose binding protein-lacZ alpha peptide
fusion protein precursor"
/protein_id="AAB86559.1"
/db_xref="GI:2623823"
/translation="MKKLLFAILVLPVFPYSHSADMKEEGLNVIWNGDKYGLARV
GKPEKDTGTVTVEHDKLEKPPQVAATCGDPDII.FWAHDRPFGVAGQLLAEIIP
DKAFQDLKPTTVDVAVNGKLIJAYPIAVEALSIIYNKDLLPNPKTWESI.PALDEL
KAKGKSALMENLOQPYETWPLJADGGVAKYKNGYDIKDVGVNAGAKAGLTFIUD
LIIKNWNAUTDYSIAAANKGETANTINGPWAWSNIDTSKVNYGVTVLPTFKQPS
KPFVGLSAGINASPNKELEFENLITDEGLEAVNKDKPLGVALKSYEEELAK
DPR.LAA.TMENAQGEIENPTIQMSAFWYAVRTAVINAASGRQTVDEALDKDAQTNSSN
NNNNNNNNIEGRISEFGSRRVDLOASIALAVLQRRDWNENPGVTLNRLAAHPF
ASWRNSEARTDRPSQQLRSLSNGEWLGCFCG"
1528..1581
/gene="malE-lacZa fusion"
/Note="M13 secretion signal sequence"
1582..2973
/gene="malB-lacZa fusion"
/product="maltose binding protein-lacZ alpha peptide
fusion protein"
3478..4338
/gene="bla"
/Note="R-plasmid ampicillin resistance gene"
3478..4338
/gene="bla"
/Note="penicillinase"
/codon_start=1
/transl_table=11
/product="RTEM-1 beta-lactamase"
/protein_id="AAB86560.1"
/db_xref="GI:2623824"
/translation="MSIQHPRVALIPPPAAFLPVPFHPETLVKVKDAEDQLGARVGY
IELDLNSGKILDSFRPEERSPMWSTPKVLICGLSVLRVADGQQLGRHRYHNSDLIV
YSPVTEKHLTDGTVTRELCSAAITMSDNTAANLLITIGPKSLTAPFLHMGDHTVRL
DRWPELNEAIPNDERDITMPVAMATTLRKLTLGELLTLASRQQLIDMEADVKAGPL
LRSLAPAGWFIADKSGAGERSGRIIAALPGDKPSRIIVYITTSQATMDERNRQIA
EIGASLIKHW"
4381..4893
/standard_name="M13 origin (+ -)"

```

ORIGIN

```

Query Match      79.7%; Score 2631; DB 12; Length 6706;
Best Local Similarity 97.7%; Pred. No. 0;
Matches 2704; Conservative 0; Mismatches 5; Indels 60; Gaps 1;

QY 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGAAGAGA 60
DB 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGAAGAGA 60

QY 61 GTCAAATTCAGGTGGTGAATGTGAACCAAGTAACTTATACGATGTCCGACAGATATCCG 120
DB 61 GTCAAATTCAGGTGGTGAATGTGAACCAAGTAACTTATACGATGTCCGACAGATATCCG 120

QY 121 GTGCTCTTTATCAGACCGTTTCCGCGTGTGAACCAAGCCAGCCAGCTTTCTGCGAAAA 180
DB 121 GTGCTCTTTATCAGACCGTTTCCGCGTGTGAACCAAGCCAGCCAGCTTTCTGCGAAAA 180

QY 181 CGCGGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTACATTCCCAACCGGTGGCAC 240
DB 181 CGCGGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTACATTCCCAACCGGTGGCAC 240

QY 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACTCCAGTCTTGCCCTGC 300
DB 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACTCCAGTCTTGCCCTGC 300

QY 301 ACSCGCGCTCGCAAAATGTTCGCGCGATTAATCTTCGCGCGATCACTGGGTGCCAGCG 360
DB 301 ACSCGCGCTCGCAAAATGTTCGCGCGATTAATCTTCGCGCGATCACTGGGTGCCAGCG 360

QY 361 TGGTGGTGTGCGATGCTAGAACGAAGCGCGCTGAAAGCTCTAAAGCGCGGTGCACAATC 420
DB 361 TGGTGGTGTGCGATGCTAGAACGAAGCGCGCTGAAAGCTCTAAAGCGCGGTGCACAATC 420

QY 421 TTCTCGCGCAACCGCTCAGTGGCTGATTAATCAATATCGCTGGATGACAGAGTGCCA 480
DB 421 TTCTCGCGCAACCGCTCAGTGGCTGATTAATCAATATCGCTGGATGACAGAGTGCCA 480

QY 481 TTGCTGTGGAAGCTGCTGCACATAATGTTCCGCGGTATTTCTTGATGTTCTTGACGAGA 540
DB 481 TTGCTGTGGAAGCTGCTGCACATAATGTTCCGCGGTATTTCTTGATGTTCTTGACGAGA 540

QY 541 CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600

QY 601 TGGTCCGATGGGTACACAGCAAAATCGCGTGTGTTAGGGGCCCATTAAGTTCCTGTCGG 660
DB 601 TGGTCCGATGGGTACACAGCAAAATCGCGTGTGTTAGGGGCCCATTAAGTTCCTGTCGG 660

QY 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720

QY 721 CGGAAACGGGAAGCGACTGGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
DB 721 CGGAAACGGGAAGCGACTGGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 780

QY 781 ATGAGGGGATCGTTCCACTCGGATGCTGTTGCCAAGCATAGATGGGCGTGGGCGCAA 840
DB 781 ATGAGGGGATCGTTCCACTCGGATGCTGTTGCCAAGCATAGATGGGCGTGGGCGCAA 840

QY 841 TGGCGGCCATTACCGAGTCCGGCTCGCGTGGTGGCGATATCTCGGTAGTGGGATACG 900
DB 841 TGGCGGCCATTACCGAGTCCGGCTCGCGTGGTGGCGATATCTCGGTAGTGGGATACG 900

QY 901 AGGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAAAACAGGATTTTC 960
DB 901 AGGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAAAACAGGATTTTC 960

QY 961 GCTGCTGGGGCAAAACAGGTGGACCGCTTGGTGGCACTCTTCAGGGCCAGGCGGTGA 1020
DB 961 GCTGCTGGGGCAAAACAGGTGGACCGCTTGGTGGCACTCTTCAGGGCCAGGCGGTGA 1020

```



```
QY 1021 AGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAAGAAAACACACCTTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAAGAAAACACACCTTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGATTCAATTAATGACGTGGCAGCAGAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGATTCAATTAATGACGTGGCAGCAGAGGTTT 1140
QY 1141 CCCGACTGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCAGAAATCTCATGTTTACAGCTTATCATGACATGCGAGGTGCACCAATGCTTCTGGG 1260
Db 1201 GCAGAAATCTCATGTTTACAGCTTATCATGACATGCGAGGTGCACCAATGCTTCTGGG 1260
QY 1261 TCAGGCGAGCATCGGAAGCTGTGGTATGCTGTGAGGTGCGTAAATCACTGCATAAATTCG 1320
Db 1261 TCAGGCGAGCATCGGAAGCTGTGGTATGCTGTGAGGTGCGTAAATCACTGCATAAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACTCCCGTCTCGGATAATGTTTTTGGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTCTCGGATAATGTTTTTGGCGCGACATCAATACGGTT 1380
QY 1381 CTGGCAAAATATTTCTGAAATGAGCTGTGACAAATTAATCATGCGTCTGTAATATGTGTGA 1440
Db 1381 CTGGCAAAATATTTCTGAAATGAGCTGTGACAAATTAATCATGCGTCTGTAATATGTGTGA 1440
QY 1441 ATTGTGAGCGGATAACAATTTACACAGAGAACCGCAGTCCGTTTTAGTGTTTTTCAGGA 1500
Db 1441 ATTGTGAGCGGATAACAATTTACACAGAGAACCGCAGTCCGTTTTAGTGTTTTTCAGGA 1500
QY 1501 GCACCTTACCAACAAAGGACCATAG----- 1524
Db 1501 GCACCTTACCAACAAAGGACCATAGCATATGAAAAATTTATTTCGCAATTCCTTTAGTG 1560
QY 1525 -----ATTATGAAAACTGAAGAGGTAACCTGGTAATCTGG 1560
Db 1561 GTACCTTTCTAFTCTCACTCGGCCGATATGAAAAATCGAAGAGGTAACCTGGTAATCTGG 1620
QY 1561 ATTTACGCGGATAAAGGCTATACCGTCTCGCTGAAGTCGGTAAGNAATTCGAGAAAGAT 1620
Db 1621 ATTTACGCGGATAAAGGCTATACCGTCTCGCTGAAGTCGGTAAGNAATTCGAGAAAGAT 1680
QY 1621 ACCGNAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCAGGTT 1680
Db 1681 ACCGNAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCAGGTT 1740
QY 1681 GCGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACGACCGCTTTGGTGCTAC 1740
Db 1741 GCGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACGACCGCTTTGGTGCTAC 1800
QY 1741 GCTCAATCTGGCCTGTGGCTGAATCAACCGGCAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1801 GCTCAATCTGGCCTGTGGCTGAATCAACCGGCAAGCGTTCCAGGACAAGCTGTAT 1860
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
Db 1861 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTGCTTACCCGATCGCTGTT 1920
QY 1861 GAAGCGTTATCGCTGATTTTAAACAAGATCTGCTCGGAAACCCGCCAAAAACCTGGGAA 1920
Db 1921 GAAGCGTTATCGCTGATTTTAAACAAGATCTGCTCGGAAACCCGCCAAAAACCTGGGAA 1980
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGAGGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1981 GAGATCCCGCGCTGGATAAAGAACTGAAGAGGTAAGAGCGCGCTGATGTTCAAC 2040
QY 1981 CTGCAAGAACCGTACTTACCTGGCGGCTGATTGCTGACGGGGTTTATCGTTCAAG 2040
Db 2041 CTGCAAGAACCGTACTTACCTGGCGGCTGATTGCTGACGGGGTTTATCGTTCAAG 2100
```

```
QY 2041 TATGAAACCGCAAGTACGACATTTAAAGACGTGGCGTGGATACGCTGGCGCGAAAGCG 2100
Db 2101 TATGAAACCGCAAGTACGACATTTAAAGACGTGGCGTGGATACGCTGGCGCGAAAGCG 2160
QY 2101 GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAACAAAACACATGAATGCAGACACCGATTAC 2160
Db 2161 GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAACAAAACACATGAATGCAGACACCGATTAC 2220
QY 2161 TCCATCGCAGAAGCTGCTTTTAATAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2221 TCCATCGCAGAAGCTGCTTTTAATAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2280
QY 2221 GCATGCTCCACATCGACACACGACAAAGTGAATTTATGSGTAAACGSETACTCGCCACCTTC 2280
Db 2281 GCATGCTCCACATCGACACACGACAAAGTGAATTTATGSGTAAACGSETACTCGCCACCTTC 2340
QY 2281 AAGGCTCAACCATCCAAACCGTTTCGTTGGCGTGTGAGCGCAGGTATTTAACGCCGCCAGT 2340
Db 2341 AAGGCTCAACCATCCAAACCGTTTCGTTGGCGTGTGAGCGCAGGTATTTAACGCCGCCAGT 2400
QY 2341 CCGAAACAAAGAGCTGGCAAAAAGAGTTCTTCGAAAACATATCTGCTGACTGATGAAGTCTG 2400
Db 2401 CCGAAACAAAGAGCTGGCAAAAAGAGTTCTTCGAAAACATATCTGCTGACTGATGAAGTCTG 2460
QY 2401 GAAGCGGTTAATAAGACAAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Db 2461 GAAGCGGTTAATAAGACAAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2520
QY 2461 TTGGGAAAGATCCACGTTATTTGCCGCCACCATGGAAAAACGCCAGAAAAGGTGAAATCATG 2520
Db 2521 TTGGGAAAGATCCACGTTATTTGCCGCCACCATGGAAAAACGCCAGAAAAGGTGAAATCATG 2580
QY 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2581 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2640
QY 2581 GCCAGCGTCTCAGACTGTGATGNAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2641 GCCAGCGTCTCAGACTGTGATGNAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2700
QY 2641 AACAAACAAACAAATAACAATAACAAACAACTCGGGATCGAGGGAAGGATTTCAGAAATTC 2700
Db 2701 AACAAACAAACAAATAACAATAACAAACAACTCGGGATCGAGGGAAGGATTTCAGAAATTC 2760
QY 2701 GGATCCTCT 2709
Db 2761 GGATCCTCT 2769
```

Search completed: August 1, 2005, 17:21:57
Job time : 9284 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 6743.2 Seconds
(without alignments)
18627.964 Million cell updates/sec

Title: US-09-765-555B-15
Perfect score: 3300
Sequence: 1 ccgacacatgaatggtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_hic:*
- 4: gb_est3:*
- 5: gb_est4:*
- 6: gb_est5:*
- 7: gb_est6:*
- 8: gb_gse1:*
- 9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	778	23.6	794	5	BQ751936	BQ751936 EST632499
2	770.6	23.4	779	7	CK394523	CK394523 hggada4D0
3	768	23.3	769	7	CK118014	CK118014 218n06.p1
4	762	23.1	793	5	BQ751655	BQ751655 EST632218
5	755.4	22.9	806	5	BQ751220	BQ751220 EST631783
6	684.8	20.8	689	6	CB863341	CB863341 HH04A08Y
7	676.4	20.5	712	1	AL038548	AL038548 DKF2p566B
8	676	20.5	710	8	AY080106	AY080106 AY080106
9	647.4	19.6	660	1	AL044483	AL044483 DKF2p434I
10	641	19.4	641	4	BJ063872	BJ063872 DKF2p3872
11	615.8	18.7	624	6	CB863814	CB863814 HH07B20Y
12	612	18.5	794	1	AL045353	AL045353 DKF2p434B
13	603	18.3	752	9	CL655575	CL655575 PRI0124A-
14	589.8	17.9	608	8	AY080094	AY080094 AY080094
15	586.4	17.8	589	1	AL037742	AL037742 DKF2p564C
16	581.2	17.6	601	1	AV594538	AV594538 AV594538
17	579	17.5	579	1	AL043868	AL043868 DKF2p434C
18	570	17.3	782	7	CK394246	CK394246 hggada1B1
19	569	17.2	774	1	AL045337	AL045337 DKF2p434A
20	563.2	17.1	585	9	CL658017	CL658017 PRI0130B
21	562.6	17.0	723	1	AL039416	AL039416 DKF2p434L
22	549.4	16.6	753	1	AL045341	AL045341 DKF2p434B
23	546.8	16.6	782	1	AL037051	AL037051 DKF2p564I
24	545.4	16.5	547	6	CA895956	CA895956 B0195F10-

C 25	545	16.5	648	7	CK394349	CK394349 hggada2D0
C 26	540.2	16.4	569	1	AL039077	AL039077 DKF2p566G
C 27	534.4	16.2	644	6	CB862524	CB862524 HH04B08U
C 28	534	16.2	734	1	AL048427	AL048427 DKF2p566H
C 29	529	16.0	740	1	AL042909	AL042909 DKF2p434J
C 30	528	16.0	720	6	CA881974	CA881974 K0994C10-
C 31	523.2	15.9	528	1	AL038811	AL038811 DKF2p566O
C 32	519.4	15.7	521	6	CA890032	CA890032 B0156F05-
C 33	519	15.7	607	6	CB862099	CB862099 HH08A02Y
C 34	516.4	15.6	518	6	CA886082	CA886082 B0124F04-
C 35	510.2	15.5	686	1	AL044407	AL044407 DKF2p434D
C 36	509.4	15.4	511	6	CA895436	CA895436 B0192A07-
C 37	498	15.1	498	1	AL039076	AL039076 DKF2p566G
C 38	494.8	15.0	1025	1	AL038025	AL038025 DKF2p566C
C 39	487.6	14.8	498	7	CK394397	CK394397 hggada2H1
C 40	477.6	14.5	491	6	CB862715	CB862715 HH03F14U
C 41	465.8	14.1	679	1	AL039128	AL039128 DKF2p566K
C 42	462.8	14.0	490	1	AL039589	AL039589 DKF2p434D
C 43	458	13.9	458	6	CA887583	CA887583 B0137H04-
C 44	456.4	13.8	471	1	AL039649	AL039649 DKF2p434G
C 45	449.2	13.6	612	6	CB863409	CB863409 HH04K05W

ALIGNMENTS

RESULT 1
BQ751936
LOCUS
DEFINITION EST632499 DSCT Colletotrichum trifolii cDNA clone pDSCT8-67, mRNA linear EST 18-JUL-2002
794 bp mRNA
ACCESSION BQ751936
VERSION BQ751936.1 GI:21907341
KEYWORDS EST.
SOURCE Colletotrichum trifolii
ORGANISM Colletotrichum trifolii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
mitosporic Phyllachoraceae; Colletotrichum.
REFERENCE 1 (bases 1 to 794)
AUTHORS Samac,D.A., Dickman,M., Town,C.D., Van Aken,S., Utterback,T.,
Cheung,F. and Fraser,C.M.
TITLE ESTs from mycelia of Colletotrichum trifolii race 1
JOURNAL Unpublished (2002)
COMMENT Other ESTs: EST632498
Contact: Deborah A. Samac
Department of Plant Pathology
University of Minnesota
495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
Tel: 612 625 1243
Fax: 651 649 5058
Email: debbyepuccini.crl.umn.edu
TIGR sequence name: MTSAH67TV More information is available at:
www.medicago.org
Seq primer: (gtA AtA CgA Ctc Act AtA ggg C).
Location/Qualifiers
1..794
/organism="Colletotrichum trifolii"
/mol_type="mRNA"
/strain="race 1"
/db_xref="taxon:5466"
/clone="pDSCT8-67"
/tissue_type="mycelia"
/inoculation="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2%glucose)."
/lab_host="DH5alpha"
/clone_lib="DSCT"

FEATURES

source
/note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2ep2 ; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda g11 from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce

E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN

Query Match 23.6%; Score 778; DB 5; Length 794;
Best Local Similarity 98.7%; Pred. No. 4.9e-219;
Matches 784; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 3 GACACCATCGAATGGTGCACAACTTCGCGGTATGCGATGATAGCGCCCGAAGAGAGT 62
Db 1 GACACTTTCGAATGGCGCAAACTTCGCGGTATGCGATGATAGCGCCCGAAGAGAGT 60

Qy 63 CAATTCAGGTTGGTGAATGTGAACACAGTAACTGATACGATGTCGAGATATGCCGGT 122
Db 61 CAATTCAGGTTGGTGAATGTGAACACAGTAACTGATACGATGTCGAGATATGCCGGT 120

Qy 123 GTCTCTTATCAGACCGTTTCGCGGTGGTGAACACAGGCGCACCGTTCCTGCGAAACG 182
Db 121 GTCTCTTATCAGACCGTTTCGCGGTGGTGAACACAGGCGCACCGTTCCTGCGAAACG 180

Qy 183 CGGGAAGAGTGAAGCGCGATGCGGAGTGAATTAATTCATTCACACCGCGTGACAA 242
Db 181 CGGGAAGAGTGAAGCGCGATGCGGAGTGAATTAATTCATTCACACCGCGTGACAA 240

Qy 243 CAACCTGCGGCAACAGTCTGCTGATGCGGTGCGACCTGCGTTCGCGGCGTGCAC 302
Db 241 CAACCTGCGGCAACAGTCTGCTGATGCGGTGCGACCTGCGTTCGCGGCGTGCAC 300

Qy 303 GCGCGCTCGCAATTTGTCGCGCGATTAATTCGCGCGATCAATTCGGGTGCGAGCGTG 362
Db 301 GCGCGCTCGCAATTTGTCGCGCGATTAATTCGCGCGATCAATTCGGGTGCGAGCGTG 360

Qy 363 GTGGTGTGATGATGAACAGCGGCTGAAGCCCTGTAAGCGGCGGTGCACATCTT 422
Db 361 GTGGTGTGATGATGAACAGCGGCTGAAGCCCTGTAAGCGGCGGTGCACATCTT 420

Qy 423 CTGCGCAACCGCTCAGTGGGCTGATCAATTAATTCGCTGCGTGGATGACCGAGTGCAT 482
Db 421 CTGCGCAACCGCTCAGTGGGCTGATCAATTAATTCGCTGCGTGGATGACCGAGTGCAT 480

Qy 483 GCTGTGGAAGTGCCTGCATTAATTTTCGCGCGTATTTCTTGATGCTCTGACCGACA 542
Db 481 GCTGTGGAAGTGCCTGCATTAATTTTCGCGCGTATTTCTTGATGCTCTGACCGACA 540

Qy 543 CCCATCAACAGTATATTTTCTCCATGAAGACGCTAGCGGCTGCGGCTGGAGCATCTG 602
Db 541 CCCATCAACAGTATATTTTCTCCATGAAGACGCTAGCGGCTGCGGCTGGAGCATCTG 600

Qy 603 GTGCGATTGGTTCACAGCAAACTCGCGCTGTAGCGGCGCCATTTAAGTCTCTGCGCG 662
Db 601 GTGCGATTGGTTCACAGCAAACTCGCGCTGTAGCGGCGCCATTTAAGTCTCTGCGCG 660

Qy 663 CGTCTGCGTCTGGCTGGCTGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
Db 661 CGTCTGCGTCTGGCTGGCTGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 720

Qy 723 GAACGGGAAGCGACTGAGTGCATGCTCGGTTTTCACAAACCATGCAATGCTGAAT 782
Db 721 GAACGGGAAGCGACTGAGTGCATGCTCGGTTTTCACAAACCATGCAATGCTGAAT 780

Qy 783 GAGGGCATCGTTC 796
Db 781 GAGGGCATCGTTC 794

RESULT 2
CK394523/c
LOCUS
DEFINITION hggda4D08 Gland Cell Amplified cDNA Library Heterodera glycines
779 bp mRNA linear EST 30-DEC-2003

cDNA, mRNA sequence.

CK394523

CK394523.1 GI:40389794

KEYWORDS

EST.

SOURCE

Heterodera glycines

Heterodera glycines

Bukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;

Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera.

1 (bases 1 to 779)

Wang,X., Allen,R., Gao,B., Goellner,M., Maier,T., Baum,T.,

Hussey,R. and Davis,E.

Submission: Wang, Allen, Gao, Goellner, Maier, Baum, Hussey, Davis

Unpublished (2003)

Contact: Tom Maier

Department of Plant Pathology, Baum Lab

Iowa State University

351 Bessey Hall, Ames, IA 50011, USA

Tel: 515-294-8854

Fax: 515-294-9420

Email: trmaier@iastate.edu

Heterodera glycines Gland Cell Amplified cDNA Library, single pass

sequence.

FEATURES

source

1..779

/organism="Heterodera glycines"

/mol_type="mRNA"

/db_xref="taxon:51029"

/cell_type="gland"

/dev_stage="mixed parasitic juvenile"

/clone_lib="Gland Cell Amplified cDNA Library"

/notes="Organ: gland cell; Vector: pSportII"

ORIGIN

Query Match 23.4%; Score 770.6; DB 7; Length 779;
Best Local Similarity 99.4%; Pred. No. 7.6e-217;
Matches 773; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 87 CCAGTAACGTTATACGATGTCGAGAGTATGCGGTGTCTCTTATCAGACCGTTTCCCGC 146
Db 779 CCAGTAACGTTATACGATGTCGAGAGTATGCGGTGTCTCTTATCAGAGTTTCCCGC 720

Qy 147 GTGGTGAACAGCGCCAGCCACCGTTTCTGCGAAACCGGGAAAAAGTGGAAAGCGCGATG 206
Db 719 GTGGTGAACAGCGCCAGCCAGCGTTTCTGCGAAACCGGGAAAAAGTGGAAAGCGCGATG 660

Qy 207 GCGGAGCTGAATTACATTCACACCGGTGGGCAACAACCTGCGGCGCAACAGTCTGTTG 266
Db 659 GCGGAGCTGAATTACATTCACACCGGTGGGCAACAACCTGCGGCGCAACAGTCTGTTG 600

Qy 267 CTGATTGGGCTTGCCACCTCCAGTCTGGCCCTGCGCGCTGCGCAAAATTTGTCGGCGG 326
Db 599 CTGATTGGGCTTGCCACCTCCAGTCTGGCCCTGCGCGCTGCGCAAAATTTGTCGGCGG 540

Qy 327 ATTAATCTCGCGCCGATCAACTGGGTGCGAGGTGGTGGTGGTGGTGGTGGTGGTGGT 386
Db 539 ATTAATCTCGCGCCGATCAACTGGGTGCGAGGTGGTGGTGGTGGTGGTGGTGGTGGT 480

Qy 387 GGGCTCGAAGCCGTAAAGCGCGGTGCAATCTTCTCGCGCAACGCGTCAGTGGGCTG 446
Db 479 GGGCTCGAAGCCGTAAAGCGCGGTGCAATCTTCTCGCGCAACGCGTCAGTGGGCTG 420

Qy 447 ATCAATTAACCTATCCGCTGGATGACCAAGATGCGCATTTGCTGTGGAAAGTGCCTGCATTA 506
Db 419 ATCAATTAACCTATCCGCTGGATGACCAAGATGCGCATTTGCTGTGGAAAGTGCCTGCATTA 360

Qy 507 GTTCCGGCGTTATTTCTTGATGTCTGTGACCGAGACCCATCAACAGATATATTTTCTCC 566
Db 359 GTTCCGGCGTTATTTCTTGATGTCTGTGACCGAGACCCATCAACAGATATATTTTCTCC 300

Qy 567 CATGAAGACGGTACCGACTGGCGGTGGAGCATCTGTCGTCATTTGGGTACCAGCAATC 626
Db 299 CATGAAGACGGTACCGACTGGCGGTGGAGCATCTGTCGTCATTTGGGTACCAGCAATC 240

Qy 627 GCGCTGTTAGCGGCCCAATTAAAGTTCTGTCTCGGGCGTCTGGCTGGCTGGCAT 686
Db 239 GCGCTGTTAGCGGCCCAATTAAAGTTCTGTCTCGGGCGTCTGGCTGGCTGGCAT 180
Qy 687 AAATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGGAGCTGGAGTGCC 746
Db 179 AAATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGGAGCTGGAGTGCC 120
Qy 747 ATGTCGGGTTTTCAACAAACCAATGCTGAATGAGGCGCATGTTTCCCACTGCGGATG 806
Db 119 ATGTCGGGTTTTCAACAAACCAATGCTGAATGAGGCGCATGTTTCCCACTGCGGATG 60
Qy 807 CTGGTTGCAACGATCAGATGGCGTGGCGGCAATGCGGCGCATTACCGAGTCCGGGC 864
Db 59 CTGGTTGCAACGATCAGATGGCGTGGCGGCAATGCGGCGCATTACCGAGTCCGGGC 2

RESULT 3
CK118014 769 bp mRNA linear EST 01-JUN-2004
LOCUS 218n06.pl AtM1 Arabidopsis thaliana cDNA clone MPWGP2011N06218
DEFINITION 5-PRIME, mRNA sequence.

ACCESSION CK118014
VERSION CK118014.1 GI:47828330
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE 1 (bases 1 to 769)
Feilner,T., Immink,R.G.H., Cahill,D.J. and Kersten,B.
AUTHORS Generation of a cDNA expression library from Arabidopsis
inflorescence meristem
TITLE Unpublished (2003)
JOURNAL Contact: Birgit Kersten
COMMENT Plant Protein Chip Group, Department Leirach
Max-Planck-Institute for Molecular Genetics
Innestr. 73, D-14195 Berlin, Germany
Tel.: +49 (0) 30/84131648
Fax: +49 (0) 30/84131128
Email: Kersten@molgen.mpg.de

Insert Length: 769 Std Error: 0.00
Plate: 218 row: N column: 6
Seq primer: pQE65.
FEATURES
Location/Qualifiers
1..769
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/ecotype="Columbia"
/db_xref="GABI:956015"
/db_xref="taxon:3702"
/clone="MPWGP2011N06218"
/tissue_type="inflorescence meristem"
/dev_stage="about one week after bolting"
/lab_host="E. coli SCS-1/pSB111"
/clone_lib="AtM1"
/note="Vector: pQE-3ONAST-attB (AY386205); Site 1: SalI;
Site 2: NotI; About 1 week after bolting, cDNA synthesis
using SuperscriptTM-system (Invitrogen) with an
oligo(dT)-primer containing NotI restriction site and a
SalI adapter. The main library (plate numbers begin with
1) of 38,000 clones was rearrayed into the sublibrary
(plate numbers begin with 201) containing 5,000 putative
expression clones. Average insert size is 1 kb. Note: The
rearrayed sublibrary (plate numbers begin with 201) was
sequenced. Library generation and sequencing was granted
in context of GABI-LAPP; data are also accessible at
<https://gabi.rzpd.de>"

ORIGIN
Query Match 23.3%; Score 768; DB 7; Length 769;
Best Local Similarity 100.0%; Pred. No. 4.5e-216;

Matches 768; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 385 GCGCGTCGAAAGCGCGGTGTAAGTCTGTCTCGCGCAACGCGTCAGTGGGC 444
Db 1 GCGCGTCGAAAGCGCGGTGTAAGTCTGTCTCGCGCAACGCGTCAGTGGGC 60
Qy 445 TGATCATTAATCAATCCGCTGGATGACAGGATGCCATTCGTGTGGAAGCTGCCTGCACTA 504
Db 61 TGATCATTAATCAATCCGCTGGATGACAGGATGCCATTCGTGTGGAAGCTGCCTGCACTA 120
Qy 505 ATGTTCCGGCGTTAATTTCTTGATGTCCTGACACAGACACCCATCAACAGATATATTTTCT 564
Db 121 ATGTTCCGGCGTTAATTTCTTGATGTCCTGACACAGACACCCATCAACAGATATATTTTCT 180
Qy 565 CCCATGAAGACGGTACCGGACTGGGCGTGAGAGCATCTGGTGCATTTGGGTCAACAGAAA 624
Db 181 CCCATGAAGACGGTACCGGACTGGGCGTGAGAGCATCTGGTGCATTTGGGTCAACAGAAA 240
Qy 625 TCGCGCTGTAGCGGGCCATTAAAGTTCTGTCTCGGGCGTCTGGCTGGCTGGCTGGC 684
Db 241 TCGCGCTGTGTAGCGGGCCATTAAAGTTCTGTCTCGGGCGTCTGGCTGGCTGGC 300
Qy 685 ATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGCGCACTGGAGTG 744
Db 301 ATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGCGCACTGGAGTG 360
Qy 745 CCATGTCGGTTTTCAACAAACCAATGCAAAATGCTGAATGAGGCGATCTTCCCACTGCGA 804
Db 361 CCATGTCGGTTTTCAACAAACCAATGCAAAATGCTGAATGAGGCGATCTTCCCACTGCGA 420
Qy 805 TGCTGGTTGCCAAGCATCAGATGCGCTGGGCGCAATGCGGCGCCATTCACGAGTCCGGGC 864
Db 421 TGCTGGTTGCCAAGCATCAGATGCGCTGGGCGCAATGCGGCGCCATTCACGAGTCCGGGC 480
Qy 865 TGCGCTGTGTGGCGATATCTCGGTAGTGGGATACGAGTACGGAACAGACAGCTCATGTT 924
Db 481 TGCGCTGTGTGGCGATATCTCGGTAGTGGGATACGAGTACGGAACAGACAGCTCATGTT 540
Qy 925 ATATCCCGCGTTAAACCAACCAATCAAAAGAGATTTTTCGCTCTGGGGCAAAACGACGTGG 984
Db 541 ATATCCCGCGTTAAACCAACCAATCAAAAGAGATTTTTCGCTCTGGGGCAAAACGACGTGG 600
Qy 985 ACCGCTGTGCTCAACTCTCTCAGGCGCGGCGGTGAAGGGCAATCAGCTGTTCGCCGTCT 1044
Db 601 ACCGCTGTGCTCAACTCTCTCAGGCGCGGCGGTGAAGGGCAATCAGCTGTTCGCCGTCT 660
Qy 1045 CACTGGTGAAAAGAAAACACCCCTTGGCGCCCAATACGCAAAACCGCTCTCCCGCGCGT 1104
Db 661 CACTGGTGAAAAGAAAACACCCCTTGGCGCCCAATACGCAAAACCGCTCTCCCGCGCGT 720
Qy 1105 TGGCCGATTCAATTAATGAGCTGGCAGACAGGTTTCCCGACTGGAAA 1152
Db 721 TGGCCGATTCAATTAATGAGCTGGCAGACAGGTTTCCCGACTGGAAA 768
RESULT 4
BO751655 793 bp mRNA linear EST 18-JUL-2002
LOCUS EST6732218 DSCT Colletotrichum trifolii cDNA clone pDSC76-59, mRNA
DEFINITION sequence.
ACCESSION BO751655
VERSION BO751655.1 GI:21907060
KEYWORDS EST.
SOURCE Colletotrichum trifolii
ORGANISM Colletotrichum trifolii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
mitosporic Phyllachoraceae; Colletotrichum.
REFERENCE 1 (bases 1 to 793)
Samac,D.A., Dickman,M., Town,C.D., Van Aken,S., Utterback,T.,
AUTHORS Cheung,F. and Fraser,C.M.
TITLE ESTs from mycelia of Colletotrichum trifolii race 1
JOURNAL Unpublished (2002)

COMMENT

Other_ESTs: EST632217
 Contact: Deborah A. Samac
 Department of Plant Pathology
 University of Minnesota
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
 Tel: 612 625 1243
 Fax: 651 649 5058
 Email: debby@puccini.crl.umn.edu
 TIGR sequence name: MTSAF59TV More information is available at:
 www.medicago.org
 Seq primer: (gtA AtA CgA Ctc Act AtA ggg C).

FEATURES

Location/Qualifiers
 1..793
 /organism="Colletotrichum trifolii"
 /mol_type="mRNA"
 /strain="race 1"
 /db_xref="taxon:5466"
 /clone="pDSC16-59"
 /tissue_type="mycelia"
 /dev_stage="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
 /lab_host="DH5alpha"
 /clone_lib="DSCT"
 /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2sp2 ; cDNA was ligated into Lambda gt11 from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN

Query Match 23.1%; Score 762; DB 5; Length 793;
 Best Local Similarity 99.1%; Pred. No. 2.8e-214;
 Matches 777; Conservative 0; Mismatches 5; Indels 2; Gaps 1;

Qy 3 GACACCATCGAATGGTGCMAAACCTTTCGGGTATGCGATGATAGCCCGGAGAGAGT 62
 Db 1 GACATCTTCGATGCGCAAAACCTTTCGGGTATGCGATGATAGCCCGGAGAGAGT 60

Qy 63 CAATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTGCGAGAGTATCGCGGT 122
 Db 61 CAATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTGCGAGAGTATCGCGGT 120

Qy 123 GTCTCTTATCAGACCGTTTCGGGTGGTGAACAGGCCAGCCACGTTTCTGCGAAAACG 182
 Db 121 GTCTCTTATCAGACCGTTTCCGGGTGGTGAACAGGCCAGCCACGTTTCTGCGAAAACG 180

Qy 183 CGGGAAAAAGTGAAGCGCGATGCGCGAGCTGAATTAACATTCACACCGCGTGGGCACAA 242
 Db 181 CGGGAAAAAGTGNAGCGCGATGCGCGAGCTGAATTAACATTCACACCGCGTGGGCACAA 240

Qy 243 CAACCTGCGGGCAACAGTGTGCTGATTTGGGTGGCCACCTCCAGTCTGGCCCTGCAC 302
 Db 241 CAACCTGCGGGCAACAGTGTGCTGATTTGGGTGGCCACCTCCAGTCTGGCCCTGCAC 300

Qy 303 GCGCGTCGCAAAATGTCGGCGGATTAATCTCGGCCCATCACTGGGTGCGACGGTG 362
 Db 301 GCGCGTCGCAAAATGTCGGCGGATTAATCTCGGCCCATCACTGGGTGCGACGGTG 360

Qy 363 GTGGTGTGATGTAGAACAGCGCGTGCAGAGCCCTGTAAAGCGCGGTGCAACAATCTT 422
 Db 361 GTGGTGTGATGTAGAACAGCGCGTGCAGAGCCCTGTAAAGCGCGGTGCAACAATCTT 420

Qy 423 CTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTCCGCTGATGACGACGATGCCATT 482
 Db 421 CTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTCCGCTGATGACGACGATGCCATT 480

Qy 483 GCTGTGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTGTGACCAGACA 542
 Db 481 GCTGTGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTGTGACCAGACA 540

Qy 543 CCCATCAACAGTATTATTTCTCCATGAAGACGGTACGGACTGGGCGTGGAGCATCTG 602
 Db 541 CCCATCAACAGTATTATTTCTCCATGAAGACGGTACGGACTGGGCGTGGAGCATCTG 600

Qy 603 GTCCATTTGGGTGACACGAGAAATCGCGCTGTAGCGGCCCATTAAGTTCTGTCTCGGCG 662
 Db 601 GTCCATTTGGGTGACACGAGAAATCGCGCTGTAGCGGCCCATTAAGTTCTGTCTCGGCG 660

Qy 663 CGTCTGCGTCTGCTGCGCATATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
 Db 661 CGTCTGCGTCTGCTGCGCATATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 720

Qy 723 GAACGGGAAGCGACTGGAGTGCATGTCGGGTTTTTCAACAAACCATGCAAAATGCTGAAT 782
 Db 721 GAACGGGAAGCGACTGGAGTGCATGTCGGGTTTTTCAACAAACCATGCAAAATGCTGAAT 778

Qy 783 GAGG 786
 Db 779 GAGG 782

RESULT 5

BO751220 806 bp mRNA linear EST 18-JUL-2002
 EST631783 DSCT Colletotrichum trifolii cDNA clone pDSC13-64, mRNA sequence.
 BO751220
 BO751220.1 GI:21906625
 EST.
 Colletotrichum trifolii
 Colletotrichum trifolii
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
 mitosporic Phyllachoraceae; Colletotrichum.
 1 (bases 1 to 806)
 Samac, D.A., Dickman, M., Town, C.D., Van Aken, S., Utterback, T.,
 Cheung, F. and Fraser, C.M.
 ESTs from mycelia of Colletotrichum trifolii race 1
 Unpublished (2002)
 Other ESTs: EST631782
 Contact: Deborah A. Samac
 Department of Plant Pathology
 University of Minnesota
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
 Tel: 612 625 1243
 Fax: 651 649 5058
 Email: debby@puccini.crl.umn.edu
 TIGR sequence name: MTSAC64TV More information is available at:
 www.medicago.org
 Seq primer: (gtA AtA CgA Ctc Act AtA ggg C).

FEATURES

Location/Qualifiers
 1..806
 /organism="Colletotrichum trifolii"
 /mol_type="mRNA"
 /strain="race 1"
 /db_xref="taxon:5466"
 /clone="pDSC13-64"
 /tissue_type="mycelia"
 /dev_stage="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
 /lab_host="DH5alpha"
 /clone_lib="DSCT"
 /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2sp2 ; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gt11 from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid

lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN		Query Match		22.9%; Score 755.4; DB 5; Length 806;		Best Local Similarity		97.4%; Pred. No. 2.5e-212;		Matches 790; Conservative 0; Mismatches 16; Indels 5; Gaps 2;	
Qy	3	GACACCATCGAATGTGTCGCAAAACCTTTTCGCGTATGGCATGATAGCGCCGGAAGAGAGT	62			1	GACACTTCGAATGCGCAAAACCTTTTCGCGTATGGCATGATAGCGCCGGAAGAGAGT	60			
Db	1	GACACTTCGAATGCGCAAAACCTTTTCGCGTATGGCATGATAGCGCCGGAAGAGAGT	60			63	CAATTACAGGGTGTGAATGTGAAACACGATTAACGATGTGCGAGATATGCCGGT	122			
Qy	63	CAATTACAGGGTGTGAATGTGAAACACGATTAACGATGTGCGAGATATGCCGGT	122			61	CAATTACAGGGTGTGAATGTGAAACACGATTAACGATGTGCGAGATATGCCGGT	120			
Db	61	CAATTACAGGGTGTGAATGTGAAACACGATTAACGATGTGCGAGATATGCCGGT	120			123	GTCTCTTATCAGACGGTTTCGCGGTGTGAACACGCGCAGCAGCTTTCTGCGAAACG	182			
Qy	123	GTCTCTTATCAGACGGTTTCGCGGTGTGAACACGCGCAGCAGCTTTCTGCGAAACG	182			121	GTCTCTTATCAGACGGTTTCGCGGTGTGAACACGCGCAGCAGCTTTCTGCGAAACG	180			
Db	121	GTCTCTTATCAGACGGTTTCGCGGTGTGAACACGCGCAGCAGCTTTCTGCGAAACG	180			183	CGGGAAAAGTGAAGCGCGCATGSCGAGCTGAATTACATTCCCAACCGGTGSCACAA	242			
Qy	183	CGGGAAAAGTGAAGCGCGCATGSCGAGCTGAATTACATTCCCAACCGGTGSCACAA	242			181	CGGGAAAAGTGAAGCGCGCATGSCGAGCTGAATTACATTCCCAACCGGTGSCACAA	240			
Db	181	CGGGAAAAGTGAAGCGCGCATGSCGAGCTGAATTACATTCCCAACCGGTGSCACAA	240			243	CAACTGGCGGCAACACAGTCGTGTGATTTGGCGTTGGCCACTCCAGTCGCGCCCTGCAC	302			
Qy	243	CAACTGGCGGCAACACAGTCGTGTGATTTGGCGTTGGCCACTCCAGTCGCGCCCTGCAC	302			241	CAACTGGCGGCAACACAGTCGTGTGATTTGGCGTTGGCCACTCCAGTCGCGCCCTGCAC	300			
Db	241	CAACTGGCGGCAACACAGTCGTGTGATTTGGCGTTGGCCACTCCAGTCGCGCCCTGCAC	300			303	CGCGCTCGCAAAATTTGTCGCGCGCATTAATATCTCGCGCGCATCAACTGGGTGCCAGCGTG	362			
Qy	303	CGCGCTCGCAAAATTTGTCGCGCGCATTAATATCTCGCGCGCATCAACTGGGTGCCAGCGTG	362			301	CGCGCTCGCAAAATTTGTCGCGCGCATTAATATCTCGCGCGCATCAACTGGGTGCCAGCGTG	360			
Db	301	CGCGCTCGCAAAATTTGTCGCGCGCATTAATATCTCGCGCGCATCAACTGGGTGCCAGCGTG	360			363	GTGGTGTGATGTAGACGAGCGCGCTCGAAGCCTGTAAGCGCGGTGCACAACTCTT	422			
Qy	363	GTGGTGTGATGTAGACGAGCGCGCTCGAAGCCTGTAAGCGCGGTGCACAACTCTT	422			361	GTGGTGTGATGTAGACGAGCGCGCTCGAAGCCTGTAAGCGCGGTGCACAACTCTT	420			
Db	361	GTGGTGTGATGTAGACGAGCGCGCTCGAAGCCTGTAAGCGCGGTGCACAACTCTT	420			423	CTCGCGCAACGCGTCAGTGGGCTGATCAATTAATCACTCCGCTGGATGACGAGATGCCATT	482			
Qy	423	CTCGCGCAACGCGTCAGTGGGCTGATCAATTAATCACTCCGCTGGATGACGAGATGCCATT	482			421	CTCGCGCAACGCGTCAGTGGGCTGATCAATTAATCACTCCGCTGGATGACGAGATGCCATT	480			
Db	421	CTCGCGCAACGCGTCAGTGGGCTGATCAATTAATCACTCCGCTGGATGACGAGATGCCATT	480			483	GCTGTGGAAGTGCCTGCACATAATGTTCCGCGTTATTTCTTGATGTCTTGACACAGACA	542			
Qy	483	GCTGTGGAAGTGCCTGCACATAATGTTCCGCGTTATTTCTTGATGTCTTGACACAGACA	542			481	GCTGTGGAAGTGCCTGCACATAATGTTCCGCGTTATTTCTTGATGTCTTGACACAGACA	540			
Db	481	GCTGTGGAAGTGCCTGCACATAATGTTCCGCGTTATTTCTTGATGTCTTGACACAGACA	540			543	CCCATCAACAGTATTTTCTCCATGAACGCTAGCGACTGGGCGTGGAGCATCTG	602			
Qy	543	CCCATCAACAGTATTTTCTCCATGAACGCTAGCGACTGGGCGTGGAGCATCTG	602			541	CCCATCAACAGTATTTTCTCCATGAACGCTAGCGACTGGGCGTGGAGCATCTG	600			
Db	541	CCCATCAACAGTATTTTCTCCATGAACGCTAGCGACTGGGCGTGGAGCATCTG	600			603	GTCCGATTGGGTTCACGCAAAATCGCGCTGTGTAGCGGCGCCCAATTAAGTTCTGTCTCGGCG	662			
Qy	603	GTCCGATTGGGTTCACGCAAAATCGCGCTGTGTAGCGGCGCCCAATTAAGTTCTGTCTCGGCG	662			601	GTCCGATTGGGTTCACGCAAAATCGCGCTGTGTAGCGGCGCCCAATTAAGTTCTGTCTCGGCG	660			
Db	601	GTCCGATTGGGTTCACGCAAAATCGCGCTGTGTAGCGGCGCCCAATTAAGTTCTGTCTCGGCG	660			663	CGTCTGCTCTGGCTGGCTGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATACG	722			
Qy	663	CGTCTGCTCTGGCTGGCTGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATACG	722			661	CGTCTGCTCTGGCTGGCTGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATACG	720			
Db	661	CGTCTGCTCTGGCTGGCTGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATACG	720			723	GAAACGGGAAGCGACTGAGTGCCATGTCGCGTTTTCACAAACCATGCAATGCTGAAT	782			
Qy	723	GAAACGGGAAGCGACTGAGTGCCATGTCGCGTTTTCACAAACCATGCAATGCTGAAT	782			721	GAAACGGGAAGCGACTGAGTGCCATGTCGCGTTTTCACAAACCATGCAATGCTGAAT	777			
Db	721	GAAACGGGAAGCGACTGAGTGCCATGTCGCGTTTTCACAAACCATGCAATGCTGAAT	777			783	GAGGCGATCGTTTCCCACTGCCATGCTGGTTG	813			
Qy	783	GAGGCGATCGTTTCCCACTGCCATGCTGGTTG	813			778	GAGGCGATCGTTTCCCACTGCCATGCTGGTTG	806			
Db	778	GAGGCGATCGTTTCCCACTGCCATGCTGGTTG	806								

RESULT 6
CB863541/c
LOCUS
DEFINITION HH04A08y HH Hordeum vulgare cDNA clone HH04A08 3-PRIME, mRNA
689 bp mRNA linear EST 22-APR-2003
sequence.

CB863541
CB863541.1 GI:30058100
EST.
KEYWORDS
SOURCE
ORGANISM
Hordeum vulgare
Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae; Hordeum.
1 (bases 1 to 689)
Varshney,R.K., Zhang,H., Burton,R., Stein,N., Langridge,P. and Graner,A.
Barley ESTs from coleoptile tissue
Unpublished (2003)
Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 689 Std Error: 0.00
Plate: 4 row: A column: 8
Seq primer: SP6.
Location/Qualifiers
1..689
/organism="Hordeum vulgare"
/mol_type="mRNA"
/cultivar="Sloop"
/db_xref="GABI:555198"
/db_xref="taxon:4513"
/clone="HH04A08"
/tissue_type="coleoptile"
/dev_stage="coleoptile, 1 day old"
/lab_host="DH10B"
/clone_lib="HH"
/note="Vector: pSPORT; Site 1: SalI (5-end of cDNA); Site 2: NotI (3-end of cDNA); Due to the cloning system used blue/white selection for recombinants is not 100 % reliable. Average insert size is 1.3 kb."

ORIGIN		Query Match		20.8%; Score 684.8; DB 6; Length 689;		Best Local Similarity		99.6%; Pred. No. 2.3e-191;		Matches 686; Conservative 0; Mismatches 3; Indels 0; Gaps 0;	
Qy	21	AAAACTTTTCGCGTATGGCATGATAGCGCCGGAAGAGAGTCAATTTCAGGTGCTGAAT	80			689	AAAACTTTTCGCGTATGGCATGATAGCGCCGGAAGAGAGTCAATTTCAGGTGCTGAAT	630			
Db	689	AAAACTTTTCGCGTATGGCATGATAGCGCCGGAAGAGAGTCAATTTCAGGTGCTGAAT	630			81	GTGAAACCAAGTAACTTATACGATGTGCGAGATATGCCGTGCTCTTATCAGACCGTT	140			
Qy	81	GTGAAACCAAGTAACTTATACGATGTGCGAGATATGCCGTGCTCTTATCAGACCGTT	140			629	GTGAAACCAAGTAACTTATACGATGTGCGAGATATGCCGTGCTCTTATCAGACCGTT	570			
Db	629	GTGAAACCAAGTAACTTATACGATGTGCGAGATATGCCGTGCTCTTATCAGACCGTT	570			141	TCCGCGTGTGTAACCAAGCGCGATGTCGGAACCGCGGGAAGAGTGAAGCG	200			
Qy	141	TCCGCGTGTGTAACCAAGCGCGATGTCGGAACCGCGGGAAGAGTGAAGCG	200			569	TCCGCGTGTGTAACCAAGCGCGATGTCGGAACCGCGGGAAGAGTGAAGCG	510			
Db	569	TCCGCGTGTGTAACCAAGCGCGATGTCGGAACCGCGGGAAGAGTGAAGCG	510			201	GCGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCACAACAACTCGCGGCAACAG	260			
Qy	201	GCGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCACAACAACTCGCGGCAACAG	260			509	GCGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCACAACAACTCGCGGCAACAG	450			
Db	509	GCGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCACAACAACTCGCGGCAACAG	450			261	TCGTTGCTGATTTGGCGTTGCCACCTTCCAGTGTGCGCCCTGCAACGCGCGCTCGCAAAATGTC	320			
Qy	261	TCGTTGCTGATTTGGCGTTGCCACCTTCCAGTGTGCGCCCTGCAACGCGCGCTCGCAAAATGTC	320			449	TGTTGCTGATTTGGCGTTGCCACCTTCCAGTGTGCGCCCTGCAACGCGCGCTCGCAAAATGTC	390			
Db	449	TGTTGCTGATTTGGCGTTGCCACCTTCCAGTGTGCGCCCTGCAACGCGCGCTCGCAAAATGTC	390			321	GCGGCGATTAATCTCGCGCGATCACTGGGTGCCAGCGTGGTGTCCGATGCTAGTA	380			
Qy	321	GCGGCGATTAATCTCGCGCGATCACTGGGTGCCAGCGTGGTGTCCGATGCTAGTA	380			389	GCGGCGATTAATCTCGCGCGATCACTGGGTGCCAGCGTGGTGTCCGATGCTAGTA	330			
Db	389	GCGGCGATTAATCTCGCGCGATCACTGGGTGCCAGCGTGGTGTCCGATGCTAGTA	330			381	CGAAGCGCGCTCGAAGCCTGTAAAGCGCGCGTGCACAATCTTCTCGCGCAACCGCTCAGT	440			
Qy	381	CGAAGCGCGCTCGAAGCCTGTAAAGCGCGCGTGCACAATCTTCTCGCGCAACCGCTCAGT	440			329	CGAAGCGCGCTCGAAGCCTGTAAAGCGCGCGTGCACAATCTTCTCGCGCAACCGCTCAGT	270			
Db	329	CGAAGCGCGCTCGAAGCCTGTAAAGCGCGCGTGCACAATCTTCTCGCGCAACCGCTCAGT	270								


```

441 GGGGTGATCATTAACATATCCGCTGGATGACAGAGATGCCATTGCTGTGGAAGCTGCCTGC 500
Db      |
269 GGGGTGATCATTAACATATCCGCTGGATGACAGAGATGCCATTGCTGTGGAAGCTGCCTGC 210
Qy      |
501 ACTAATGTTCCGGCGTTATTCTTGTGATGCTCTGACAGACACCCCATCAACAGTATTATT 560
Db      |
209 ACTAATGTTCCGGCGTTATTCTTGTGATGCTCTGACAGACACCCCATCAACAGTATTATT 150
Qy      |
561 TTCTCCCATGAAGACGGTACGCGACTGGCGTGGAGCATCTGGTCGCAATTTGGSTCACAG 620
Db      |
149 TTCTCCCATGAAGACGGTACGCGACTGGCGTGGAGCATCTGGTCGCAATTTGGSTCACAG 90
Qy      |
621 CAAATCGCGCTGTAGCGGCCCAATTAAGTCTGTCGCGCGTCTGCGTCTGGCTGCG 680
Db      |
89 CAAATCGCGCTGTAGCGGCCCAATTAAGTCTGTCGCGCGTCTGCGTCTGGCTGCG 30
Qy      |
681 TGGCATAAATATCTCACTCGCAATCAAT 709
Db      |
29 TGGCATAAATATCTCACTCGCAATCAAT 1

RESULT 7
AL038548/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AL038548
DKFZp566E0346.r1.566 (synonym: hfkd2) Homo sapiens cDNA clone
DKFZp566E0346.5, mRNA sequence.
AL038548
AL038548.1 GI:5407738
EST.
Homo sapiens
Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 712)
Ottenwaelder,B., Obermaier,B., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
EST (Ottenwaelder, et al.)
Unpublished (1995)
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by MediGenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project. No s1 sequence
available.
This clone (DKFZp566E0346) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1. 712
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp566E0346"
/tissue_type="kidney"
/dev_stage="fetal"
/lab_host="Xi-2blue"
/clone_lib="566 (synonym: hfkd2)"
/notes="Vector: pAMPI; Site_1: NotI; Site_2: SalI"

FEATURES
source
Query Match 20.5%; Score 676.4; DB 1; Length 712;
Best Local Similarity 99.6%; Pred. No. 7.2e-189;
Matches 699; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
Qy 33 GGTATGCATGATAGCGCCCGGAGAGAGTCAATTCAGGGTGGTGAATGTGAACCAAGTA 92
Db 712 GGTATGCATGATAGCGCCCGGAGAGAGTCAATTCAGGGTGGTGAATGTGAACCAAGTA 653
Qy 93 ACCTTATACCATGTCGAGAGTATCGGGTGTCTCTTATCAGACCGCTTCCCGGTGGTG 152

```

```

Db      |
Qy      |
592 AACAGGCGCAGCCAGCTTTCTGCGAAACCGGGG-AAAAAGTGAAGCGGCGATGGCGGA 211
592 AACAGGCGCAGCCAGCTTTCTGCGAAACCGGGGAAAAAGTGAAGCGGCGATGGCGGA 533
212 GCTGAATTAACATTCGCAACCGCGTGGCAACAACCTGCGCGGCAACAGTGTGCTGAT 271
532 GCTGAATTAACATTCGCAACCGCGTGGCAACAACCTGCGCGGCAACAGTGTGCTGAT 473
272 TGGCGTTGCGCACTCCAGTCTGGCCCTGCAC-GGCGCGTGCCTGCAATTTGTGCGCGCATTA 330
472 TGGCGTTGCGCACTCCAGTCTGGCCCTGCACGGCGCGCTGCAATTTGTGCGCGCATTA 413
331 AATCTCGCGCGCATCAACTGGGTGCCAGCGTGGTGTGTCGATGCTAGAGAACGAGCGCG 390
412 AATCTCGCGCGCATCAACTGGGTGCCAGCGTGGTGTGTCGATGCTAGAGAACGAGCGCG 353
391 TCGAAGCCTGTAAAGCGCGGTGCACAACTCTTCTCGCGCAACCGCTCAGTGGCTGATCA 450
352 TCGAAGCCTGTAAAGCGCGGTGCACAACTCTTCTCGCGCAACCGCTCAGTGGCTGATCA 293
451 TTAACATATCGCTGGATGACAGGATGCCATTGCTGTGGAAGCTGCTGCACATAATGTTTC 510
292 TTAACATATCGCTGGATGACAGGATGCCATTGCTGTGGAAGCTGCTGCACATAATGTTTC 233
511 CGCGGTATTCTTGATGCTCTGACAGACACCCATCAACAGTATTATTTCTCCCATG 570
232 CGCGGTATTCTTGATGCTCTGACAGACACCCATCAACAGTATTATTTCTCCCATG 173
571 AAGACGGTACGGCACTGGGCGTGGAGCATCTGTCGATGGTGGTGCACCAACATCGCG 630
172 AAGACGGTACGGCACTGGGCGTGGAGCATCTGTCGATGGTGGTGCACCAACATCGCG 113
631 TGTAGCGGCGCCATTAAAGTCTGTCGCGCGCTGTCGCTGCTGGTGGCTGGCATAAAT 690
112 TGTAGCGGCGCCATTAAAGTCTGTCGCGCGCTGTCGCTGCTGGTGGCTGGCATAAAT 53
691 ATCTCACTCGCATCAATCAATTCAGCGATAGCGGAACGGGAAG 732
52 ATCTCACTCGCATCAATCAATTCAGCGATAGCGGAACGGGAAG 11

AV080106 710 bp DNA linear GSS 06-NOV-2002
AY080106 Scripps Pier (La Jolla, CA) uncultured virus community
uncultured marine virus genomic clone SIO51p3G6L, genomic survey
sequence.
AY080106
AY080106.1 GI:24745302
GSS.
uncultured marine virus
uncultured marine virus
Viruses; environmental samples.
1 (bases 1 to 710)
Breitbart,M., Salamon,P., Andresen,B., Mahaffy,J.M., Segall,A.M.,
Mead,D., Azam,F. and Rohwer,F.
Genomic analysis of uncultured marine viral communities
Proc. Natl. Acad. Sci. U.S.A. 99 (22), 14250-14255 (2002)
22294988
12384570
Contact: Rohwer F
Biology Dept.
San Diego State University
5500 Campanile Dr, San Diego, CA 92102, USA
Tel: 6195941336
Fax: 619595676
Email: forest@sunstroke.sdsu.edu
Class: shotgun.
Location/Qualifiers
1. 710
/organism="uncultured marine virus"

FEATURES
source

```

/mol_type="genomic DNA"
/db_xref="taxon:186617"
/clone="SI051p3G6L"
/clone_lib="Scripps Pier (La Jolla, CA) uncultured virus community"
/note="Marine viruses were isolated from 200 liters of surface seawater using a combination of differential filtration and density-dependent gradient centrifugation. Linker-amplified shotgun libraries were constructed by randomly shearing the total marine viral community DNA, end-repairing, ligating dsDNA linkers to the ends, and amplifying the fragments using Vent DNA polymerase. The resulting fragments were ligated into the pSMART vector and electroporated into MC12 cells (Lucigen; Middleton, WI)"

ORIGIN

Query Match 20.5%; Score 676; DB 8; Length 710;
Best Local Similarity 98.9%; Pred. No. 9.5e-189; Indels 1; Gaps 1;
Matches 690; Conservative 0; Mismatches 7;

Qy 400 GTAAAGCGCGGTGCACAACTCTTCGCGCAACGCGTCAGTGGCTGATCACTAACTATC 459
Db |||||
8 GTGCTGCAGCGGTGCACAACTCTTCGCGCAACGCGTCAGTGGCTGATCACTAACTATC 67
Qy 460 CGCTGGATGACAGGATGCCATTGCTGTGGAAGCTGCTGCACATAATGTTCCGGCGTTAT 519
Db |||||
68 CGCTGGATGACAGGATGCCATTGCTGTGGAAGCTGCTGCACATAATGTTCCGGCGTTAT 127
Qy 520 TTCTTGATGTCCTGACAGACACCCATCAACAGTATTATTTCTCCCATGAAGCGGTA 579
Db |||||
128 TTCTTGATGTCCTGACAGACACCCATCAACAGTATTATTTCTCCCATGAAGCGGTA 187
Qy 580 CGCAGCTGGCGGTGAGCATCTGTGCGATTGGGTCAACAGCAAACTCGCGTGTAGCGG 639
Db |||||
188 CGGACTGGCGGTGAGCATCTGTGCGATTGGGTCAACAGCAAACTCGCGTGTAGCGG 247
Qy 640 GCCATTAAAGTTCTGCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTC 699
Db |||||
248 GCCATTAAAGTTCTGCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTC 307
Qy 700 GCAATCAATTCAGCCGATAGCGGAACGGGAGCGACTGCTCCCACTGCGATGCTGCTGTCGCAACG 819
Db |||||
368 AACAAACCATGCAATGCTGAATGAGGCACTGTTCCCACTGCGATGCTGTTGCCAAG 427
Qy 820 ATCAGATGGCGTGGCGCAATGCGCGCATTPACCGAGTTCGGGCTGCGCGTGTGTCGG 879
Db |||||
428 ATCAGATGGCGTGGCGCAATGCGCGCATTPACCGAGTTCGGGCTGCGCGTGTGTCGG 487
Qy 880 ATATCTCGGTAGTGGATACGACGATACGAGACAGCTCATGTTATATCCCGCGTTAA 939
Db |||||
488 ATATCTCGGTAGTGGATACGACGATACGAGACAGCTCATGTTATATCCCGCGTTAA 547
Qy 940 CCACCATCAACAGGATTTTCGCTGCTGGGCAACACGAGTGGACCGCTTGTGTCAC 999
Db |||||
548 CCACCATCAACAGGATTTTCGCTGCTGGGCAACACGAGTGGACCGCTTGTGTCAC 607
Qy 1000 TCTCTCAGGGCGAGCGGTGAAGGCAATCAAGCTGTGTCGCGCTCACTGTTGAAAGAA 1059
Db |||||
608 TCTCTCAGGGCGAGCGGTGAAGGCAATCAAGCTGTGTCGCGCTCACTGTTGAAAGAA-AA 666
Qy 1060 AAACACCTGCGGCCCAATAGCGAAACCGCTCTCCC 1097
Db |||||
667 AAACACCTGCGGCCCAATAGCGAAACCGNCTTTCCC 704

RESULT 9
AL044483/c

LOCUS AL044483 660 bp mRNA linear EST 04-SEP-2003

DEFINITION DKFZp4341102_s1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION DKFZp4341102.3, mRNA sequence.
VERSION AL044483
KEYWORDS AL044483.1 GI:5432701
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 660)
AUTHORS Ansoorge,W., Benes,V., Krieger,S., Mewes,H.W., Gassenhuber,J. and Wismann,S.
TITLE EST (Ansoorge, Benes, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 3' sequence of the clone insert
Clone from S. Wismann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wismann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No r1 sequence available.
This clone (DKFZp4341102) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
Location/Qualifiers
1..660
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp4341102"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 19.6%; Score 647.4; DB 1; Length 660;
Best Local Similarity 99.7%; Pred. No. 2.9e-180;
Matches 659; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 85 AACAGTAACCTTATACGATGTCGAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCC 144
Db |||||
660 AACAGTAACCTTATACGATGTCGAGAGTATGCCGGTGTCTCTTATCAGACCGTTTCC 601
Qy 145 CGTGGTGAACAGCGCCAGCCAGCTTCTTCGGAACACCGGGGAAAAAGTGGAGCGGGA 204
Db |||||
600 CGTGGTGAACAGCGCCAGCCAGCTTCTTCGGAACACCGGGGAAAAAGTGGAGCGGGA 541
Qy 205 TGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAACAACCTGGCGGGCAACAGTCT 264
Db |||||
540 TGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAACAACCTGGCGGGCAACAGTCT 481
Qy 265 TGCTGATTGGGTTGCCACCTCCAGTCTGGCCCTGCACCGCGCTCGCAAAATTTGCGGG 324
Db |||||
480 TGCTGATTGGGTTGCCACCTCCAGTCTGGCCCTGCACCGCGCTCGCAAAATTTGCGGG 421
Qy 325 CGATTAAATCTCGCGCCGATCAACTGGGTGCAGCGTGGTGTGATGGTAGAACGAA 384
Db |||||
420 CGATTAAATCTCGCGCCGATCAACTGGGTGCAGCGTGGTGTGATGGTAGAACGAA 361
Qy 385 GCGGCTCGAAGCTGTAAAGCGGGTGCACAACTTCTCGCAACCGCTCAGTGGGC 444
Db |||||
360 GCGGCTCGAAGCTGTAAAGCGGGTGCACAACTTCTCGCAACCGGTGAGTGGGC 301
Qy 445 TGATCATTAATATCCGCTGGATGACACGAGTGCATTGCTGTGGAAGCTGCTGCACATA 504
Db |||||
300 TGATCATTAATATCCGCTGGATGACACGAGTGCATTGCTGTGGAAGCTGCTGCACATA 241
Qy 505 ATGTTCCGGCGTTATTTCTTGATGTCTCTGACGACACCCATCAACAGTATTTTCT 564

```

Db      240  ATGTTCCGGGTTATTTCTTGATGTTCTCTGACACAGACACCCATCAACAGTATTATTTCT 181
QY      565  CCATGAAGACGGTACGCGACTGGCGGTGAGCATCTGTCGATTTGGGTCAACAGCAAA 624
Db      180  CCATGAAGACGGTACGCGACTGGCGGTGAGCATCTGTCGATTTGGGTCAACAGCAAA 121
QY      625  TCGCGCTGTTAGCGGGGCCCAATTAAGTTCTGTCGCGGGGTCTGCGTCTGGCTGGCTGGC 684
Db      120  TCGCGCTGTTAGCGGGGCCCAATTAAGTTCTGTCGCG-GCGTCTGGCTGGCTGGC 62
QY      685  ATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAAGCGGAGCGACTGGAGTG 744
Db      61  ATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAAGCGGAGCGACTGGAGTG 2
QY      745  C 745
Db      1  C 1

RESULT 10
BJ063872/c
LOCUS
DEFINITION
BJ063872 NIBB Mochii normalized Xenopus tailbud library Xenopus
laevis cDNA clone XL077k02 5', mRNA sequence.
ACCESSION
BJ063872
VERSION
BJ063872.1 GI:17471074
SOURCE
Xenopus laevis (African clawed frog)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Xenopus.
1 (bases 1 to 641)
Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i.T. and
Kohara,Y.
Expressed genes in X. laevis embryo
Unpublished (2001)
Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp
The information of this clone is available through the following
URL.
http://xenopus.nibb.ac.jp.
Location/Qualifiers
1..641
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="XL077k02"
/tissue_type="whole embryo"
/dev stage="stage 25"
/clone.lib="NIBB Mochii normalized Xenopus tailbud
library"

FEATURES
source
Query Match 19.4%; Score 641; DB 4; Length 641;
Best Local Similarity 100.0%; Pred. No. 2.3e-178;
Matches 641; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      30  CGCGGTATGCATAGCCCGCGAGAGAGTCAATTCAGGGTGTGATGTGAACCA 89
Db      641  CGCGGTATGCATAGCCCGCGAGAGAGTCAATTCAGGGTGTGATGTGAACCA 582
QY      90  GTAAACGTTATACGATGTGCGAGAGTATGCCGGGTGTCCTTATCAGACCGTTCCCGCGTG 149
Db      581  GTAAACGTTATACGATGTGCGAGAGTATGCCGGGTGTCCTTATCAGACCGTTCCCGCGTG 522
QY      150  GTGAACCGCCAGCCACGTTTCTGCGAAACCGCGGAAAAAGTGGAGCGCGATGGCG 209

```

```

Db      521  GTGAACCGCCAGCCACGTTTCTGCGAAACCGGGGAAAAGTGGAAAGCGCGCATGGCG 462
QY      210  GAGCTGAATTACATTCACACCGCGTGGGCAACAACATGGCGGGCAACAGTCTGTTGCTG 269
Db      461  GAGCTGAATTACATTCACACCGCGTGGGCAACAACATGGCGGGCAACAGTCTGTTGCTG 402
QY      270  ATTGGCGTTGGCACTCCAGTCTCGCCCTGCAACGCGCGGTGGCAAAATGTCGGGCGCAT 329
Db      401  ATTGGCGTTGGCACTCCAGTCTCGCCCTGCAACGCGCGGTGGCAAAATGTCGGGCGCAT 342
QY      330  AAATCTCGCGCGCATCAACTGGGTGCGAGCGTGTGTCGATGGTAGAACGAAGCGGC 389
Db      341  AAATCTCGCGCGCATCAACTGGGTGCGAGCGTGTGTCGATGGTAGAACGAAGCGGC 282
QY      390  GTCGAAGCGCTGTAAGCGCGGTGCAAACTTCTCGCGCAACGCGTCACTGGGCTGATC 449
Db      281  GTCGAAGCGCTGTAAGCGCGGTGCAAACTTCTCGCGCAACGCGTCACTGGGCTGATC 222
QY      450  ATTAATCTATCCGTGGATGACAGGATGCCATTTGCTGTGGAAGCTGCCTGCACATAATGTT 509
Db      221  ATTAATCTATCCGTGGATGACAGGATGCCATTTGCTGTGGAAGCTGCCTGCACATAATGTT 162
QY      510  CCGCGCTTATTTCTTGATGTCTCTGACCAAGACACCATCAACAGTATTTTCTCCCAT 569
Db      161  CCGCGCTTATTTCTTGATGTCTCTGACCAAGACACCATCAACAGTATTTTCTCCCAT 102
QY      570  GAAGACGGTACGCGACTGGGCGTGGAGCATCTGGTCGCATTTGGTCCAGCAACATCGCG 629
Db      101  GAAGACGGTACGCGACTGGGCGTGGAGCATCTGGTCGCATTTGGTCCAGCAACATCGCG 42
QY      630  CTGTTAGCGCGCCATTAAGTTCTGTCGCGCGGCTCGC 670
Db      41  CTGTTAGCGCGCCATTAAGTTCTGTCGCGCGGCTCGC 1

CB863814 624 bp mRNA linear EST 22-APR-2003
HH07B20Y HH Hordeum vulgare cDNA clone HH07B20 3-PRIME, mRNA
sequence.
CB863814
ACCESSION
CB863814.1 GI:30058373
VERSION
CB863814.1
KEYWORDS
EST.
SOURCE
Hordeum vulgare
ORGANISM
Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
1 (bases 1 to 624)
Varshney,R.K., Zhang,H., Burton,R., Stein,N., Langridge,P. and
Graner,A.
Barley ESTs from coleoptile tissue
Unpublished (2003)
Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 624 Std Error: 0.00
Plate: 7 row: B column: 20
Seq primer: SP6.
Location/Qualifiers
1...624
/organism="Hordeum vulgare"
/mol_type="mRNA"
/cultivar="Sloop"
/db_xref="GABI:555471"
/db_xref="taxon:4513"
/clone="HH07B20"
/tissue_type="coleoptile"

```

ORIGIN	/dev_stage="coleoptile, 1 day old"	
	/lab_host="DH10B"	
TITLE	EST (Duesterhoeft, et al.)	
	Unpublished (1999)	
JOURNAL	Contact: MIPS	
	MIPS	
COMMENT	Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany	
	This is the 3' sequence of the clone insert	
EST (Duesterhoeft, et al.)	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer	
	Research Center (DKFZ); sequenced by Qiagen within the cDNA	
Unpublished (1999)	sequencing consortium of the German Genome Project. r1 sequence	
	also available.	
Contact: MIPS	This clone is available at the RZPD in Berlin.	
	Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059	
MIPS	Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.	
	Location/Qualifiers	
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany	1. .794	
	/organism="Homo sapiens"	
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer	/mol_type="mRNA"	
	/db_xref="taxon:9606"	
Research Center (DKFZ); sequenced by Qiagen within the cDNA	/clone="DKFZp434B075"	
	/tissue_type="testis"	
sequencing consortium of the German Genome Project. r1 sequence	/dev_stage="adult"	
	/lab_host="DH10B"	
also available.	/clone_lib="434 (synonym: htes3)"	
	/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"	
This clone is available at the RZPD in Berlin.	ORIGIN	
	Query Match 18.5%; Score 612; DB 1; Length 794;	
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059	Best Local Similarity 99.7%; Pred. NO. 1e-169;	
	Matches 634; Conservative 0; Mismatches 0; Indels 2; Gaps 2;	
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.	QY 569 TGAAGCGGTACGCGACTGGCGGTGGAGCATCTGGTCGATTTGGTACCGCAATCGC 628	
	Db 794 TGAAGCGGTACGCGACTGGCGGTGGAG-ATCTGGTCGATTTGGTACCGCAATCGC 736	
Location/Qualifiers	QY 629 GCTGTTAGCGGCGCCATTAAAGTTCTGCTCGCGCGTCTGGTCTGCTGGCTGGCTGGCATAA 688	
	Db 735 GCTGTTAGCGGCGCCATTAAAGTTCTGCTCGCGCGTCTGGTCTGCTGGCTGGCTGGCATAA 677	
1. .794	QY 689 ATATCTCACTCGCAATCAAAATTCAGCCGATACGCGAACGGGAAGCGCATGAGAGTGCAT 748	
	Db 676 ATATCTCACTCGCAATCAAAATTCAGCCGATACGCGAACGGGAAGCGCATGAGAGTGCAT 617	
/organism="Homo sapiens"	QY 749 GTCCGGTTTTCAACAAACCATGCAATGCTGAATGAGGGCATCGTCCCATCGCATGCT 808	
	Db 616 GTCCGGTTTTCAACAAACCATGCAATGCTGAATGAGGGCATCGTCCCATCGCATGCT 557	
/mol_type="mRNA"	QY 809 GGTTCGCAACGATCAGATGGCGCTGGCGCAATGCGGCCATTACCGAGTCCGGCTGCG 868	
	Db 556 GGTTCGCAACGATCAGATGGCGCTGGCGCAATGCGGCCATTACCGAGTCCGGCTGCG 497	
/db_xref="taxon:9606"	QY 869 GGTTCGCGGATATCTCGGTAGTGGGATACGAGATACCGAAGACAGCTCATGTTATAT 928	
	Db 496 GGTTCGCGGATATCTCGGTAGTGGGATACGAGATACCGAAGACAGCTCATGTTATAT 437	
/clone="DKFZp434B075"	QY 929 CCCGCGTTTAAACCATCAAAACAGGATTTTCGCTGCTGGGGCAAAACAGCGTGGACCG 988	
	Db 436 CCCGCGTTTAAACCATCAAAACAGGATTTTCGCTGCTGGGGCAAAACAGCGTGGACCG 377	
/tissue_type="testis"	QY 989 CTTCGTCAACTCTCTCAGGCGCAGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACT 1048	
	Db 376 CTTCGTCAACTCTCTCAGGCGCAGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACT 317	
/dev_stage="adult"	QY 1049 GGTGAAAAAGAAAAACCACTCCGCGCCCAATACGCAAAACCGCTCTCCCGCGCGTTGGC 1108	
	Db 316 GGTGAAAAAGAAAAACCACTCCGCGCCCAATACGCAAAACCGCTCTCCCGCGCGTTGGC 257	
/lab_host="DH10B"	QY 1109 CGATTCAATATGAGCTGGCAAGAGGTTTCCGACTGGAAACGGGCGAGTAGCGCA 1168	
	Db 256 CGATTCAATATGAGCTGGCAAGAGGTTTCCGACTGGAAACGGGCGAGTAGCGCA 197	
/clone_lib="434 (synonym: htes3)"	QY 1169 ACGCAATTAATGTAGTTAGCTCACTCAATTAGGCAC 1204	
	Db 196 ACGCAATTAATGTAGTTAGCTCACTCAATTAGGCAC 161	
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"	RESULT 12	
	AL045353/c	
Query Match 18.7%; Score 615.8; DB 6; Length 624;	LOCUS	
	DEFINITION	
Best Local Similarity 99.7%; Pred. NO. 7e-171;	DKFZp434B075_e1 434 (synonym: htes3) Homo sapiens cDNA clone	
	Matches 617; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
Mismatches 617; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	QY 3 GACACCATCGAATGTCGAAACCTTTTCGCGGTATGGCATGATAGCGCCGCGAAGAGAGT 62	
	Db 619 GACACCATCGAATGTCGCGCAACCTTTTCGCGGTATGGCATGATAGCGCCGCGAAGAGAGT 560	
/dev_stage="coleoptile, 1 day old"	QY 63 CAATTACAGGTGGTGAATGTGAACACAGTAAAGTATACGATGTCGAGAGTATCCCGGT 122	
	Db 559 CAATTACAGGTGGTGAATGTGAACACAGTAAAGTATACGATGTCGAGAGTATCCCGGT 500	
/lab_host="DH10B"	QY 123 GTCTCTTATCAGACCGTTTCCGCGTGTGTGAACACGAGCCAGCTTTCTCGGAAAAACG 182	
	Db 499 GTCTCTTATCAGACCGTTTCCGCGTGTGTGAACACGAGCCAGCTTTCTCGGAAAAACG 440	
Score 615.8; DB 6; Length 624;	QY 183 CGGGAAAAAGTGGAAAGCGCGGATGCGGAGCTGAATTAATTCCTCCAAACCGGTGGCAAA 242	
	Db 439 CGGGAAAAAGTGGAAAGCGCGGATGCGGAGCTGAATTAATTCCTCCAAACCGGTGGCAAA 380	
Pred. NO. 7e-171;	QY 243 CAATTCGCGGCGCAACAGTCTGTTGCTGATTTGCGCTTGCACCTCCAGTCTGSCCTCGCAC 302	
	Db 379 CAATTCGCGGCGCAACAGTCTGTTGCTGATTTGCGCTTGCACCTCCAGTCTGSCCTCGCAC 320	
Mismatches 2; Indels 0; Gaps 0;	QY 303 GCGCGCTGCGCAATTTGTCGCGCGATTAATCTCGCGCGCATCAATCGGGTGCCAGCGTG 362	
	Db 319 GCGCGCTGCGCAATTTGTCGCGCGATTAATCTCGCGCGCATCAATCGGGTGCCAGCGTG 260	
Indels 0; Gaps 0;	QY 363 GTGGTGTGATGTGATAGACGAGCGGCTGAGAGCTGTAAAGCGGCGGTGCACAACTTT 422	
	Db 259 GTGGTGTGATGTGATAGACGAGCGGCTGAGAGCTGTAAAGCGGCGGTGCACAACTTT 200	
Length 624;	QY 423 CTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACCAAGTGCCTATT 482	
	Db 199 CTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACCAAGTGCCTATT 140	
Pred. NO. 7e-171;	QY 483 GCTGTGGAAGCTGCTGCACTAATGTTCCGGGTTATTTCTTGATGTCTTGACAGACA 542	
	Db 139 GCTGTGGAAGCTGCTGCACTAATGTTCCGGGTTATTTCTTGATGTCTTGACAGACA 80	
Mismatches 2; Indels 0; Gaps 0;	QY 543 CCCATCAACAGTATTTTCTCCATGAAGACGTAACGCGACTGGGCGTGGAGCATCTG 602	
	Db 79 CCCATCAACAGTATTTTCTCCATGAAGACGTAACGCGACTGGGCGTGGAGCATCTG 20	
Score 615.8; DB 6; Length 624;	QY 603 GTCGCAATTTGGTTCACCGC 621	
	Db 19 GTCGCAATTTGGTTCACCGC 1	
Pred. NO. 7e-171;	RESULT 12	
	AL045353/c	
Mismatches 2; Indels 0; Gaps 0;	LOCUS	
	DEFINITION	
Indels 0; Gaps 0;	DKFZp434B075_e1 434 (synonym: htes3) Homo sapiens cDNA clone	
	Matches 617; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
Accession	QY 621	
	Db 621	
Version	QY 621	
	Db 621	
Keywords	QY 621	
	Db 621	
Source	QY 621	
	Db 621	
Organism	QY 621	
	Db 621	
Reference	QY 621	
	Db 621	
Authors	QY 621	
	Db 621	

RESULT 13
CL655575/c
LOCUS
DEFINITION
CL655575 752 bp DNA linear GSS 09-JUL-2004
PRI0124a_C01 - PRI0124a.B21 (752) Mixed stage fosmid library of P.
pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.
CL655575
CL655575.1 GI:50134897
GSS.
Pristionchus pacificus
Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
1 (bases 1 to 752)
Srinivasan, J., Otto G.W., Kahlow, U., Geisler, R. and Sommer, R.J.
AppaDB: an AcceDB database for the nematode satellite organism
Pristionchus pacificus
Nucleic Acids Res. 32 (1), D421-D422 (2004)
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: raif.sommer@tuebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end
sequenced at Vancouver, Canada.
Seq primer: T7
Class: fosmid ends.
Location/Qualifiers
1..752
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage fosmid library of P. pacificus
var. California"
/note="Vector: pBpifos-5 Fosmid vector"

ORIGIN

Query Match 18.3%; Score 603; DB 9; Length 752;
Best Local Similarity 98.9%; Pred. No. 4.7e-167;
Matches 628; Conservative 0; Mismatches 5; Indels 2; Gaps 2;
Qy 1996 TTCACCTGGCCGCTGATGTGCTGAC - GGCGGTTATGCGTTCAAGTATGAAACGGCAA 2054
Db 752 TTCACCTGGCCGCTGATGTGCTGACGGGGGTTATGCGTTCAAGTATGAAACGGCAA 693
Qy 2055 GTACGACATTAAGACGTTGGCGTGGATACGTTGGCGGAAACGGGTTCTGACCTTCTCT 2114
Db 692 GTACGACATTAAGACGTTGGCGTGGATACGTTGGCGGAAACGGGTTCTGACCTTCTCT 633
Qy 2115 GGTGACCTGATTAAGGCGAAACAGCGATGACCATCAACGGCCGCTGGGTCATGTTCCAAAT 2174
Db 632 GGTGACCTGATTAAGGCGAAACAGCGATGACCATCAACGGCCGCTGGGTCATGTTCCAAAT 573
Qy 2175 TGCCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCGCTGGGTCATGTTCCAAAT 2234
Db 572 TGCCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCGCTGGGTCATGTTCCAAAT 513
Qy 2235 CGACACGACAAAGTGAATTTATGTTGTTAAACGGTACTCGGACCTTCAAGGTTCAACCATC 2294
Db 512 CGACACGACAAAGTGAATTTATGTTGTTAAACGGTACTCGGACCTTCAAGGTTCAACCATC 453
Qy 2295 CAACACCTGTTGGCTGTCTGACGCGAGGTATTAACGGCCGAGTCCGAAACAGAGCT 2354
Db 452 CAACACCTGTTGGCTGTCTGACGCGAGGTATTAACGGCCGAGTCCGAAACAGAGCT 393
Qy 2355 GCGAAAGAGTTCTCGAAACCTATCTGCTGATGAAGGTCTGGAAGCGGTTAA-TA 2413
Db 392 GCGAAAGAGTTCTCGAAACCTATCTGCTGATGAAGGTCTGGAAGCGGTTAA-TA 333

Qy 2414 AAGACAAACCGCTGGTGCCCTAGCGTGAAGTCTTACGAGGAAGAGTTGCGGAAAGATC 2473
Db 332 AGGCCAAACCGCTGGTGCCCTAGCGTGAAGTCTTACGAGGAAGAGTTGCGGAAAGATC 273
Qy 2474 CACGTATTGCGCGCCACCATGGAACCCCAAGAGTGAATCATGCCGAACATCCCGC 2533
Db 272 CACGTATTGCGCGCCACCATGGAACCCCAAGAGTGAATCATGCCGAACATCCCGC 213
Qy 2534 AGATGTCGCGTTTCTGCTGATGCGCTGCTACTGCGGTGATCAACGCCCGCAGCGGTGTC 2593
Db 212 AGATGTCGCGTTTCTGCTGATGCGCTGCTACTGCGGTGATCAACGCCCGCAGCGGTGTC 153
Qy 2594 AGATGTCGATGAAGCCCTGAAGACGCGCAGACT 2628
Db 152 AGACTGTCGATGAAGCCCTGAAGACGCGCAGACT 118
RESULT 14
AY080094/c
LOCUS
DEFINITION
AY080094 608 bp DNA linear GSS 06-NOV-2002
uncultured marine virus genomic clone SIO51p3D4L, genomic survey
sequence.
AY080094
ACCESSION
AY080094.1 GI:24745276
VERSION
GSS.
KEYWORDS
uncultured marine virus
SOURCE
uncultured marine virus
ORGANISM
Viruses; environmental samples.
REFERENCE
1 (bases 1 to 608)
Breitbart, M., Salamon, P., Andresen, B., Mahaffy, J.M., Segall, A.M.,
Mead, D., Azam, F. and Rohwer, F.
Genomic analysis of uncultured marine viral communities
Proc. Natl. Acad. Sci. U.S.A. 99 (22), 14250-14255 (2002)
22294988
MEDLINE
12384570
PUBMED
COMMENT
Contact: Rohwer F
Biology Dept.
San Diego State University
5500 Campanile Dr, San Diego, CA 92102, USA
Tel: 6195941336
Fax: 619595676
Email: forest@unstroke.sdsu.edu
Class: shotgun.
Location/Qualifiers
1..608
/organism="uncultured marine virus"
/mol_type="genomic DNA"
/db_xref="taxon:186617"
/clone="SIO51p3D4L"
/clone_lib="Scripps Pier (La Jolla, CA) uncultured virus
community"
/note="Marine viruses were isolated from 200 liters of
surface seawater using a combination of differential
filtration and density-dependent gradient centrifugation.
Linker-amplified shotgun libraries were constructed by
randomly shearing the total marine viral community DNA,
end-repairing, ligating dsDNA linkers to the ends, and
amplifying the fragments using Vent DNA polymerase. The
resulting fragments were ligated into the pSMART vector
and electroporated into MC12 cells (Lucigen; Middleton,
WI)."

ORIGIN

Query Match 17.9%; Score 589.8; DB 8; Length 608;
Best Local Similarity 99.7%; Pred. No. 3.7e-163;
Matches 591; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 287 CAGTCTGGCCCTGACGCGCGCTCCCAATTCGCGGCGATTAATCTCGCGCGATCA 346
Db 608 CAGTCTGGCCCTGACGCGCGCTCCCAATTCGCGGCGATTAATCTCGCGCGATCA 549

QY 347 ACTGGTCCACGCTGCTGCTGATGGTAGAAGAGCGCGTCGAAGCTGTAAAGC 406
Db ACTGGTCCACGCTGCTGCTGATGGTAGAAGAGCGCGTCGAAGCTGTAAAGC 489
QY 407 GCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCAATAACTATCCGCTGGA 466
Db GCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCAATAACTATCCGCTGGA 429
QY 467 TGACGAGATGCATGCTGTGGAAAGCTGCGTGCACATAATGTTCCAGCGTTATTTCTTGA 526
Db TGACGAGATGCATGCTGTGGAAAGCTGCGTGCACATAATGTTCCAGCGTTATTTCTTGA 369
QY 527 TGTCTCTGACGAGACACCCATCAACAGTATATTTCTCCCATGAAGCGGTACCGGACT 586
Db TGTCTCTGACGAGACACCCATCAACAGTATATTTCTCCCATGAAGCGGTACCGGACT 309
QY 587 GGGCGTGGAGCATCTGGTTCGCAATGGGTCAACAGTATATTTCTCCCATGAAGCGGTACCGGACT 646
Db GGGCGTGGAGCATCTGGTTCGCAATGGGTCAACAGTATATTTCTCCCATGAAGCGGTACCGGACT 249
QY 647 AAGTTCTGTCTGCGCGCGTCTGCGTCTGGCTGGGCAATAATATCTACCTCGCAATCA 706
Db AAGTTCTGTCTGCGCGCGTCTGCGTCTGGCTGGGCAATAATATCTACCTCGCAATCA 189
QY 707 AATTGAGCGATAGCGGACGGAAGGCGACTGGAGTGCATGTCGGTTTTCACAAAC 766
Db AATTGAGCGATAGCGGACGGAAGGCGACTGGAGTGCATGTCGGTTTTCACAAAC 129
QY 767 CATGCAATGCTGAATGAGGCGATGTTCCCACTCCGATGCTGGTTGCCAAGATCAGAT 826
Db CATGCAATGCTGAATGAGGCGATGTTCCCACTCCGATGCTGGTTGCCAAGATCAGAT 69
QY 827 GCGCTGGCGCAATGCGCGCATTAACCGAGTCCGGGCTGCGCGTGGTGCGG 879
Db GCGCTGGCGCAATGCGCGCATTAACCGAGTCCGGGCTGCGCGTGGTGCGG 16

RESULT 15
AL037742/c
LOCUS DKFZp564C157.r1.564 (synonym: hfbr2) Homo sapiens cDNA clone
DEFINITION DKFZp564C157.5', mRNA sequence.
ACCESSION AL037742
VERSION 1 (bases 1 to 589)
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 589)
Bloeker, H., Boecker, M., Brandt, P., Mewes, H.W., Gassenhuber, J. and
Wiemann, S.
EST (Bloeker, et al.)
Unpublished (1999)
CONTACT: MIPS
MIPS

Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wilmann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wilmann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.

No sl sequence available.
This clone (DKFZp564C157) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers

1..589
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp564C157"

FEATURES
source

/tissue_type="brain"
/dev_stages="fetal"
/lab_host="X1-2blue"
/clone_lib="564 (synonym: hfbr2)"
/note="Vector: pAMP1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 17.8%; Score 586.4; DB 1; Length 589;
Best Local Similarity 99.7%; Pred. No. 3.7e-162;
Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 517 TATTTCTTGATGCTCTGACGAGACACCCATCAACAGTATATTTCTCCCATGAAGACG 576
Db TATTTCTTGATGCTCTGACGAGACACCCATCAACAGTATATTTCTCCCATGAAGACG 530
QY 577 GTACGCGATCGGGCGTGGAGCATCTGGTTCGATTTGGGTCAACAGCAAAATCGCGCTGTAG 636
Db GTACGCGATCGGGCGTGGAGCATCTGGTTCGATTTGGGTCAACAGCAAAATCGCGCTGTAG 470
QY 637 CGGCGCCATTAAGTCTGCTCGGGCGCTCTGCGTCTGGCTGGTGGCATAAATATCTCA 696
Db CGGCGCCATTAAGTCTGCTCGGGCGCTCTGCGTCTGGCTGGTGGCATAAATATCTCA 410
QY 697 CTCGCAATCAAAATTCAGCCGATAGCGGAACGGAAGGCGACTGGAGTCCCATGTCGGTT 756
Db CTCGCAATCAAAATTCAGCCGATAGCGGAACGGAAGGCGACTGGAGTCCCATGTCGGTT 350
QY 757 TTCAACAAACCATGCAATGCTGAATGAGGCGATCGTTCCCACTCCGATGCTGGTTGCCA 816
Db TTCAACAAACCATGCAATGCTGAATGAGGCGATCGTTCCCACTCCGATGCTGGTTGCCA 290
QY 817 ACGATCAGATGGCGCTGGGCGCAATGCGCGCAATTAACGAGTCCGGGCTGCGCGTTGGTG 876
Db ACGATCAGATGGCGCTGGGCGCAATGCGCGCAATTAACGAGTCCGGGCTGCGCGTTGGTG 230
QY 877 CGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCGT 936
Db CGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCGT 170
QY 937 TAACCAACCATCAAAACAGGATTTTTCGCTGCTGGGCGCAACACAGCGTGGACCGCTTGTGTC 996
Db TAACCAACCATCAAAACAGGATTTTTCGCTGCTGGGCGCAACACAGCGTGGACCGCTTGTGTC 110
QY 997 AACTCTCTCAGGGCGAGCGCGTGAAGGGCAATCAGCTGTTGCCCGCTCTCACTGGTGA AAA 1056
Db AACTCTCTCAGGGCGAGCGCGTGAAGGGCAATCAGCTGTTGCCCGCTCTCACTGGTGA AAA 50
QY 1057. GAAAAACCAACCTCGCGCCCAATACGCAACCGCCTCTCCCGCGCGTT 1105
Db GAAAAACCAACCTCGCGCCCAATACGCAACCGCCTCTCCCGCGCGTT 1

Search completed: August 3, 2005, 02:02:48
Job time : 6745.2 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 1309.2 Seconds
(without alignments)
16314.373 Million cell updates/sec

Title: US-09-765-555B-15
Perfect score: 3300
Sequence: 1 ccgacaccatcgatgtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3300	100.0	3300	10	US-09-765-555-14
2	3300	100.0	3300	10	US-09-765-555-15
3	3226.4	97.8	3300	10	US-09-765-555-18
4	3215.2	97.4	3300	10	US-09-765-555-17
5	3207.2	97.2	3300	10	US-09-765-555-16
6	2705.8	82.0	7373	18	US-10-257-384A-1
7	2705.8	82.0	8101	17	US-10-149-472-5

8	2701	81.8	6648	19	US-10-343-859-8	Sequence 8, Appli
9	2698	81.8	9191	19	US-10-343-859-9	Sequence 9, Appli
10	2690.4	81.5	7112	18	US-10-263-153-40	Sequence 40, Appli
11	2690.4	81.5	7259	18	US-10-263-153-30	Sequence 30, Appli
12	2690.4	81.5	7322	18	US-10-263-153-35	Sequence 35, Appli
13	2690.4	81.5	7352	18	US-10-263-153-25	Sequence 25, Appli
14	2690.4	81.5	7370	18	US-10-263-153-20	Sequence 20, Appli
15	2690.4	81.5	7370	18	US-10-263-153-61	Sequence 61, Appli
16	2690.4	81.5	7370	18	US-10-263-153-66	Sequence 66, Appli
17	2690.4	81.5	7370	18	US-10-263-153-71	Sequence 71, Appli
18	2690.4	81.5	7403	18	US-10-263-153-15	Sequence 15, Appli
19	2690.4	81.5	7442	18	US-10-263-153-10	Sequence 10, Appli
20	2690.4	81.5	7478	18	US-10-263-153-3	Sequence 3, Appli
21	2636.8	79.9	6806	10	US-09-908-943A-194	Sequence 194, App
22	2636.8	79.9	6806	20	US-10-801-487-194	Sequence 194, App
23	2636.8	79.9	6806	20	US-10-801-938-194	Sequence 194, App
24	2636.8	79.9	6806	20	US-10-801-509-194	Sequence 194, App
25	2636.8	79.9	6806	20	US-10-801-486-194	Sequence 194, App
26	2636.8	79.9	6806	21	US-10-801-493-194	Sequence 194, App
27	2605.4	79.0	7553	18	US-10-263-153-7	Sequence 7, Appli
28	1609.4	48.8	5558	16	US-10-241-596-137	Sequence 137, App
29	1474	44.7	4700	17	US-10-313-963A-54	Sequence 54, Appli
30	1462.4	44.3	4921	21	US-10-495-491-2	Sequence 2, Appli
31	1462.4	44.3	4935	21	US-10-495-491-1	Sequence 1, Appli
32	1462.4	44.3	4945	21	US-10-495-491-3	Sequence 3, Appli
33	1462.4	44.3	4951	21	US-10-495-491-4	Sequence 4, Appli
34	1417.6	43.0	1511	10	US-09-987-763-30	Sequence 30, Appli
35	1417.6	43.0	5926	13	US-10-024-809-3	Sequence 3, Appli
36	1238.4	37.5	5024	17	US-10-313-963A-52	Sequence 52, Appli
37	1238.4	37.5	5448	21	US-10-671-995A-2	Sequence 2, Appli
38	1238.4	37.5	5454	21	US-10-671-995A-1	Sequence 1, Appli
39	1237.2	37.5	1922	10	US-09-995-898A-32	Sequence 32, Appli
40	1237.2	37.5	1922	17	US-10-420-034A-32	Sequence 32, Appli
41	1237.2	37.5	6270	21	US-10-855-897-2	Sequence 2, Appli
42	1204	36.5	4557	17	US-10-457-372-5	GENERAL INFORMA
43	1204	36.5	6664	18	US-10-270-176-1	Sequence 1, Appli
44	1204	36.5	6668	18	US-10-270-176-2	Sequence 2, Appli
45	1204	36.5	6742	18	US-10-270-176-19	Sequence 19, Appli

ALIGNMENTS

RESULT 1
US-09-765-555-14
; Sequence 14, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765.555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m1 and zinc finger
; OTHER INFORMATION: protein ZfPm1
US-09-765-555-14

Query Match 100.0%; Score 3300; DB 10; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTCAAAAACTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTCAAAAACTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTACAGGTGGTGAATGTGAACACAGTAACGTTATACCATGTGCGAGAGTATGCG 120
Db 61 GTCAATTACAGGTGGTGAATGTGAACACAGTAACGTTATACCATGTGCGAGAGTATGCG 120
QY 121 GTGTCTTATCAGACCGTGTTCGCCGGTGGTGAACACAGGCCAGCCACGCTTTCTGGAAAA 180
Db 121 GTGTCTTATCAGACCGTGTTCGCCGGTGGTGAACACAGGCCAGCCACGCTTTCTGGAAAA 180
QY 181 CCGGGGAAAAAGTGGAGCGGCGATGGCGAGCTGAATTAACATTCCTCCAACCGCGTGGCAC 240
Db 181 CCGGGGAAAAAGTGGAGCGGCGATGGCGAGCTGAATTAACATTCCTCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGCAACAGTCGTGCTGATTTGGGGTGGTCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAACAGTCGTGCTGATTTGGGGTGGTCCACCTCCAGTCTGGCCCTGC 300
QY 301 AGCGCCCGTCGCAAAATGTTCGGCGGATTAATCTTCGGCGCGATCAACTGGGTGCCAGCG 360
Db 301 AGCGCCCGTCGCAAAATGTTCGGCGGATTAATCTTCGGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTTCGATGCGTAAGCAAGACCGCGCTCGAAGCCTGTAAAGCGCGCGTGCACAATC 420
Db 361 TGGTGGTTCGATGCGTAAGCAAGACCGCGCTCGAAGCCTGTAAAGCGCGCGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTTCGGCGCGATCAACTGGGTGCCAGCG 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTTCGGCGCGATCAACTGGGTGCCAGCG 480
QY 481 TTGTGTGGAGTGCCTGCACTAATGTTCCGGGTTATTTCTTGATGTTCTCTGACCCAGA 540
Db 481 TTGTGTGGAGTGCCTGCACTAATGTTCCGGGTTATTTCTTGATGTTCTCTGACCCAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGCTACCGCATGGCGGTGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGCTACCGCATGGCGGTGAGCATC 600
QY 601 TGGTGGCATGGGTCCACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCTCTCTCGG 660
Db 601 TGGTGGCATGGGTCCACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCTCTCTCGG 660
QY 661 CGGCTCTGGCTGGCTGGCATAAATCTCACTTCGCAATCAAAATCAGCCGATAG 720
Db 661 CGGCTCTGGCTGGCTGGCATAAATCTCACTTCGCAATCAAAATCAGCCGATAG 720
QY 721 CCGAACCGGAAGGCGACTGGAGTGCCATGTCGGGTTTCAACCAACCATGCAATGCTGA 780
Db 721 CCGAACCGGAAGGCGACTGGAGTGCCATGTCGGGTTTCAACCAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATCGTTCACCTGCGATGCTGGTTCGAACGATCAGATGGCGCTGGGCGCAA 840
Db 781 ATGAGGGCATCGTTCACCTGCGATGCTGGTTCGAACGATCAGATGGCGCTGGGCGCAA 840
QY 841 TGCGCGCATTAACGAGTCCGGCTGGCGTGGTGGCGGATATCTCGGTAGTGGGATACG 900
Db 841 TGCGCGCATTAACGAGTCCGGCTGGCGTGGTGGCGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACCCATCAACACAGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACCCATCAACACAGATTTTC 960
QY 961 GCTGTGGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCTGTGTGGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTTCCCGTCTCACTGTTGAAGAAAAACACCTCGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTTCCCGTCTCACTGTTGAAGAAAAACACCTCGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGCGTGGCCGATTAATTAATGCAGTGGCACGACAGGTTT 1140

Db 1081 CGCAAAACCGCTCTCCCGCGCGTGGCCGATTAATTAATGCAGTGGCACGACAGGTTT 1140
QY 1141 CCCGACTGGAAAGCGGCGAGTGAAGCGCAACGCAATAATATGTAGTTAGCTCACTCATTAG 1200
Db 1141 CCCGACTGGAAAGCGGCGAGTGAAGCGCAACGCAATAATATGTAGTTAGCTCACTCATTAG 1200
QY 1201 GCACAAATCTCATGTTTTGACAGCTTATCATCGATGCAAGTGCACCGTGACCAATGCTTCGGCG 1260
Db 1201 GCACAAATCTCATGTTTTGACAGCTTATCATCGATGCAAGTGCACCGTGACCAATGCTTCGGCG 1260
QY 1261 TCAGGACAGCCATCGEAAAGCTGTGATGCTGTGACGCTGTAATACCTGCAATAATTCG 1320
Db 1261 TCAGGACAGCCATCGEAAAGCTGTGATGCTGTGACGCTGTAATACCTGCAATAATTCG 1320
QY 1321 TGTGCTCAAGGGCGACCTCCCGTTCGGATAATGTTTTTTCGCGCGACATCAATAACGGTT 1380
Db 1321 TGTGCTCAAGGGCGACCTCCCGTTCGGATAATGTTTTTTCGCGCGACATCAATAACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGTGGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGTGGA 1440
QY 1441 ATTGTGAGCGGATAACAAATTTTCACAGGAAACAGCCAGTCCGTTTAAAGTGTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTTTCACAGGAAACAGCCAGTCCGTTTAAAGTGTTCACGA 1500
QY 1501 GCATTTCCACCAACAGGACCATAGATTGAAACTGGAAGAGTAATTCGGTAATCTCG 1560
Db 1501 GCATTTCCACCAACAGGACCATAGATTGAAACTGGAAGAGTAATTCGGTAATCTCG 1560
QY 1561 ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGATCACCGTTGAGCATTCGGATAAACTTGAAGAGAAATTCGCCACAGTT 1680
Db 1621 ACCGGAATTAAGATCACCGTTGAGCATTCGGATAAACTTGAAGAGAAATTCGCCACAGTT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTCGGGCAACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTCGGGCAACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAGCTGTAT 1800
QY 1801 CGGTTTACCTGGGATGCGGTAACGTTACACGCGCAAGCTGATTGCTTACCGGATCGCTGTT 1860
Db 1801 CGGTTTACCTGGGATGCGGTAACGTTACACGCGCAAGCTGATTGCTTACCGGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTTATACAAAGATCTGCTGCCGAACCGCGCAAAACCTGGGA 1920
Db 1861 GAAGCGTTATCGCTGATTTTATACAAAGATCTGCTGCCGAACCGCGCAAAACCTGGGA 1920
QY 1921 GAGATCCCGCGCTGATATAAGAACTGAAAGCGAAAGTAAAGAGCGCGCTGTATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGATATAAGAACTGAAAGCGAAAGTAAAGAGCGCGCTGTATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTTCACCTGGCGCTGATTGCTGTGACGGGGTTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGGCGCTGATTGCTGTGACGGGGTTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGACGTTGGGCGTGATTAACGCTGGCGGAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGACGTTGGGCGTGATTAACGCTGGCGGAAGCG 2100
QY 2101 GGTCTGACCTTCTGTTGACCTGATTAATAACACATCAATCAATGACAGACCCGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTGACCTGATTAATAACACATCAATGACAGACCCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCCCTTTTAAAGCGGAAACAGCGATGACCATCAACCGCCGCTGG 2220

```
Db 2161 TCCATCGCAGAGCTGCCTTTTAAAGGCGAAGAACAGCGATGACCATCAACGGGCCGTGG 2220
Qy 2221 GCATGGTCAACATCGACACCAAGAAAGTGAATTTATGGTGAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACCAAGAAAGTGAATTTATGGTGAACGGTACTGCGGACCTTC 2280
Qy 2281 AAGGCTCAACCATCAAAACCGTTCTGTTGGCGTGTGAGGCGAGGTATTAAACGCCGCCAGT 2340
Db 2281 AAGGCTCAACCATCAAAACCGTTCTGTTGGCGTGTGAGGCGAGGTATTAAACGCCGCCAGT 2340
Qy 2341 CGAAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACCTATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CGAAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACCTATCTGCTGACTGATGAAGGTCTG 2400
Qy 2401 GAAGCGGTAAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGAAAG 2460
Db 2401 GAAGCGGTAAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGAAAG 2460
Qy 2461 TTGGCGAAAGATCCACGTTATTCGCCGCCACCATGGAAAGCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATTCGCCGCCACCATGGAAAGCGCCAGAAAGGTGAATCATG 2520
Qy 2521 CGAAACATCCGCGAGATGTCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CGAAACATCCGCGAGATGTCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGGTGCTGACAGTGTGATGAAGCCCTGAAAGAGCGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGCTGACAGTGTGATGAAGCCCTGAAAGAGCGCGCAGACTAAATTCGAGCTCG 2640
Qy 2641 AACACACAAACAAATAACAAACAAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACACACAAACAAATAACAAACAAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Qy 2701 GGATCCTCTTCTCTGTGCGCCAGCGGCCCTCGAGCCGCGGGAAGGCCCTATGCTGT 2760
Db 2701 GGATCCTCTTCTCTGTGCGCCAGCGGCCCTCGAGCCGCGGGAAGGCCCTATGCTGT 2760
Qy 2761 CCGGAATGTGTAAGTCTTCTCTCAGAGCTCTCACCTGGTGGCGCACACAGCGTACCCAC 2820
Db 2761 CCGGAATGTGTAAGTCTTCTCTCAGAGCTCTCACCTGGTGGCGCACACAGCGTACCCAC 2820
Qy 2821 ACGGTTGAAAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTAGCGAGTCCAGCAAC 2880
Db 2821 ACGGTTGAAAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTAGCGAGTCCAGCAAC 2880
Qy 2881 CTGTTGCGCATCAACGCACTCATCTGCGAGAGAGCCATACAAATGTCAGAAATGTGCG 2940
Db 2881 CTGTTGCGCATCAACGCACTCATCTGCGAGAGAGCCATACAAATGTCAGAAATGTGCG 2940
Qy 2941 AAGTCTTTCTCTCGGTCTGACAACTCTGTCGGCACCAACGTAATCAACACCGGGGAGAG 3000
Db 2941 AAGTCTTTCTCTCGGTCTGACAACTCTGTCGGCACCAACGTAATCAACACCGGGGAGAG 3000
Qy 3001 CCTATGCTTGTCCGAATGTGGTAAGTCTTTCAGCGCAGCGATAAATCTGTTGGCGCCAC 3060
Db 3001 CCTATGCTTGTCCGAATGTGGTAAGTCTTTCAGCGCAGCGATAAATCTGTTGGCGCCAC 3060
Qy 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTAGC 3120
Qy 3121 CAGGCGGGCACCTGGCGAGCCATCAACGCACTCTACTGCGGAGAGGCAATACAAATGT 3180
Db 3121 CAGGCGGGCACCTGGCGAGCCATCAACGCACTCTACTGCGGAGAGGCAATACAAATGT 3180
Qy 3181 CCAGAAATGCGCAAGTCTTTCTCTCGGTCTGCAATCTCGTCCGCGACCAACGTAATCAC 3240
Db 3181 CCAGAAATGCGCAAGTCTTTCTCTCGGTCTGCAATCTCGTCCGCGACCAACGTAATCAC 3240
Qy 3241 ACCGGTAAAAAACTAGTGGCCAGGCGCGCGAGTACCGGTACGAGTTCGGGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCCAGGCGCGCGAGTACCGGTACGAGTTCGGGACTACGCT 3300
```

RESULT 2

```
US-09-765-555-15
; Sequence 15, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m2 and zinc finger
; OTHER INFORMATION: protein ZfPm2
US-09-765-555-15
```

Query Match 100.0%; Score 3300; DB 10; Length 3300;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0; Matches 3300; Conservative 0;

```
Qy 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGCGGTATGGCATGATAGCCCGGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGCGGTATGGCATGATAGCCCGGGAAGAGA 60
Qy 61 GTCAATTGAGGTGGTGAATGTG AAAACCGAGTAAAGTTATACGATGTCGAGAGTATGCGG 120
Db 61 GTCAATTGAGGTGGTGAATGTG AAAACCGAGTAAAGTTATACGATGTCGAGAGTATGCGG 120
Qy 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAAACCGAGCGAGCCAGCCAGTTCCTCGGAAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAAACCGAGCGAGCCAGCCAGTTCCTCGGAAAA 180
Qy 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 240
Qy 241 AACAACTGGCGGGCAAAACAGTCTGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGCAAAACAGTCTGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGGCGCTGCGCAATTCGCGCGGATTAATCTCGCGCGATTAATCTCGCGCGATTAATCTCGCGCG 360
Db 301 ACGGCGCTGCGCAATTCGCGCGGATTAATCTCGCGCGATTAATCTCGCGCGATTAATCTCGCGCG 360
Qy 361 TGGTGGTCTCGATGGTAGAAGCGCGCTCGAAGCCTGTAAAGCGGGGTGCACAATC 420
Db 361 TGGTGGTCTCGATGGTAGAAGCGCGCTCGAAGCCTGTAAAGCGGGGTGCACAATC 420
Qy 421 TTCTCGCGCAACCGGTGATGGGTGATCATTAATCTCGGTGATGACCAAGGATGCCA 480
Db 421 TTCTCGCGCAACCGGTGATGGGTGATCATTAATCTCGGTGATGACCAAGGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCCCTGACATTAATGTTCCGGCGTTATTTCTTGATGTTCTTGATGTTCT 540
Db 481 TTGCTGTGGAAGCTGCCCTGACATTAATGTTCCGGCGTTATTTCTTGATGTTCTTGATGTTCT 540
Qy 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACGCGACTGGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACGCGACTGGGCGGTGGAGCATC 600
Qy 601 TGGTCCGATTTGGGTCCACGACAAATTCGCGCTGTTAGCGGGGCCCATTAAGATTCTGTCTGG 660
```

Db 601 TGGTCGCATTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATAAGTTCTGTCTCGG 660
Qy 661 CGCGTCTCGCTGCTGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTGCTGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAAACGGAAAGCGACATGGAGTGCATGTCCGGTTTTCAACAAACCAATGCAAAATGCTGA 780
Db 721 CGGAAACGGAAAGCGACATGGAGTGCATGTCCGGTTTTCAACAAACCAATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATCGTTCCCACTCGGATGCTGGTTGCCAAGATCGAGTGGCGTGGCGCAA 840
Db 781 ATGAGGGCATCGTTCCCACTCGGATGCTGGTTGCCAAGATCGAGTGGCGTGGCGCAA 840
Qy 841 TGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACCAACATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACCAACATCAAAACAGGATTTTC 960
Qy 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTTGCCGCTCTCACTGTTGAAAGAAACCAACCTGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCGCTCTCACTGTTGAAAGAAACCAACCTGCGGCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCATTAAATGCAGCTGGCAGCAGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCATTAAATGCAGCTGGCAGCAGAGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGGCAACCGCAATTAATGTAGTTAGCTCACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGGCAACCGCAATTAATGTAGTTAGCTCACTCATTAG 1200
Qy 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACAGGTGCACCAATGCTTCTGCGG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACAGGTGCACCAATGCTTCTGCGG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGTATGGTGTGCGAGTGTAAATCACTGCATAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGTATGGTGTGCGAGTGTAAATCACTGCATAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTTGGCCGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTTGGCCGACATCATACGGTT 1380
Qy 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
Qy 1441 ATTGTGAGCGGATAACAAATTTCAACAGGAAACAGCCAGTTCAGTGTGTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTTCAACAGGAAACAGCCAGTTCAGTGTGTTTTCACGA 1500
Qy 1501 GCACCTTCAACCAACAGGACCATAGATTATGAAACTGAAAGGTAAACTGGTAAATCTGG 1560
Db 1501 GCACCTTCAACCAACAGGACCATAGATTATGAAACTGAAAGGTAAACTGGTAAATCTGG 1560
Qy 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAACTCGGTAAAGAAATTCGAGAAGAT 1620
Db 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAACTCGGTAAAGAAATTCGAGAAGAT 1620
Qy 1621 ACCGGAATTAAGTACCGTTGAGCTCCGGAATAAATCGGAAGAAATTCACACAGGTT 1680
Db 1621 ACCGGAATTAAGTACCGTTGAGCTCCGGAATAAATCGGAAGAAATTCACACAGGTT 1680
Qy 1681 CGCGCAATCGCGATGGCCCTGCATTTATCTCTGGGCACACGACCGCTTTGGTGCTAC 1740

Db 1681 CGCGCAATCGCGATGGCCCTGCATTTATCTTCTGGGCACACGACCGCTTTGGTGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGTGTGAATCAACCCCGGACAAAGCGTTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGTGTGAATCAACCCCGGACAAAGCGTTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTTACCTGGGATGCGGTACGTTAAACGGCAAGCTGATGCTTACCAGATCGCTGTT 1860
Db 1801 CCGTTTTACCTGGGATGCGGTACGTTAAACGGCAAGCTGATGCTTACCAGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAACCTCGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAACCTCGGAA 1920
Qy 1921 GAGATCCCGGCTGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGTATGTTCAAC 1980
Db 1921 GAGATCCCGGCTGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGTATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTTCACCTGCGCTGATTTGCTGACCGGGGTTATGCTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGCGCTGATTTGCTGACCGGGGTTATGCTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTTAAAGACGTGGCGTGGATAAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTTAAAGACGTGGCGTGGATAAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGTTGACCTGATTTAAAAACAAACATGATGAGCAACCGGATTC 2160
Db 2101 GGTCTGACCTTCTGTTGACCTGATTTAAAAACAAACATGATGAGCAACCGGATTC 2160
Qy 2161 TCCATCGCAAGCTGCTTTTAAAGCGAAACAGCGATGACCATCAACGCCCGGTGG 2220
Db 2161 TCCATCGCAAGCTGCTTTTAAAGCGAAACAGCGATGACCATCAACGCCCGGTGG 2220
Qy 2221 GCATGTGCCAAATCAACACACAGCAAAAGTGAATTTATGGTGTAAACGGTACTGCCGACCTTC 2280
Db 2221 GCATGTGCCAAATCAACACACAGCAAAAGTGAATTTATGGTGTAAACGGTACTGCCGACCTTC 2280
Qy 2281 AAGGTTCAACCATCCAAACCGTTGCTGGGTGCTGAGCGCAGGTATTAACGCCGCGAGT 2340
Db 2281 AAGGTTCAACCATCCAAACCGTTGCTGGGTGCTGAGCGCAGGTATTAACGCCGCGAGT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAAAGATTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGATTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGTATGAGCTGAAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGTATGAGCTGAAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGGCGAAAGATCCAGTATTCGCCACCATGGAATAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTCGCCACCATGGAATAACGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAAACATCCCGCAGATGTCGGTTTCTGGTATGCGCGTCTGCTGCTGATGCTGCTGCTG 2580
Db 2521 CCGAAACATCCCGCAGATGTCGGTTTCTGGTATGCGCGTCTGCTGCTGATGCTGCTGCTG 2580
Qy 2581 GCCAGCGTCTGAGACTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGAGACTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACACAAACAAATAACAAATAACAACTCCGGATCGAGGAGGATTTTCAGAAATTC 2700
Db 2641 AACACAAACAAATAACAAATAACAACTCCGGATCGAGGAGGATTTTCAGAAATTC 2700
Qy 2701 GGATCTCTTCTCTGTGGCCAGCGGCTTCGAGCCCGGGGAGAAAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTGGCCAGCGGCTTCGAGCCCGGGGAGAAAGCCCTATGCTTGT 2760
Qy 2761 CCGGATGTGGTAAGTCTTCTCTCAGAGCTCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2820
Db 2761 CCGGATGTGGTAAGTCTTCTCTCAGAGCTCTCAGCTGCTCAGCTGCTGCTGCTGCTGCTGCTG 2820

```
QY 2821 ACCGGTGAACCGGTATATAATGCCAGAGTGGCGCAAACTTTTAGCCAGTCCAGCAAC 2880
D 2821 ACCGGTGAACCGGTATATAATGCCAGAGTGGCGCAAACTTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGGTGGCCCACTCAACGCACCTCATCTGGCGAAGGCCATACAAATGTCCAGATGTGGC 2940
D 2881 CTGGTGGCCCACTCAACGCACCTCATCTGGCGAAGGCCATACAAATGTCCAGATGTGGC 2940
QY 2941 AAGTCTTTTCTCTCGGTCTGACAACTCTGTCGGCACCAACGTAFTCAACCGGGGAGAAG 3000
D 2941 AAGTCTTTTCTCTCGGTCTGACAACTCTGTCGGCACCAACGTAFTCAACCGGGGAGAAG 3000
QY 3001 CCCTATGCTTGTCCGGAAATGGTAAAGTCTTTCAGCCGCGAGGATAAAGTGGTGGCCAC 3060
D 3001 CCCTATGCTTGTCCGGAAATGGTAAAGTCTTTCAGCCGCGAGGATAAAGTGGTGGCCAC 3060
QY 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAATGCCAGAGTGGCGCAAACTTTTAGC 3120
D 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAATGCCAGAGTGGCGCAAACTTTTAGC 3120
QY 3121 CAGCGCGGCCACCTGGCCAGCCATCAACGCACCTCATCTGGCGAAGGCCATACAAATGT 3180
D 3121 CAGCGCGGCCACCTGGCCAGCCATCAACGCACCTCATCTGGCGAAGGCCATACAAATGT 3180
QY 3181 CCAGAAATGTGGCAAGTCTTTCTCGGTCTGACAACTCTGTCGGCACCAAGCTACTCAC 3240
D 3181 CCAGAAATGTGGCAAGTCTTTCTCTCGGTCTGACAACTCTGTCGGCACCAAGCTACTCAC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCCAGCGCCGAGTACCGGTACGAGCTTCCGAGCTACGCT 3300
D 3241 ACCGGTAAAAAACTAGTGGCCAGCGCCGAGTACCGGTACGAGCTTCCGAGCTACGCT 3300
```

RESULT 3

```
US-09-765-555-18
; Sequence 18, Application US/0976555
; Publication No. US2003003735SA1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-Ap3 and zinc finger
; OTHER INFORMATION: protein 2FPAP3
US-09-765-555-18
```

```
Query Match 97.8%; Score 3226.4; DB 10; Length 3300;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3254; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
```

```
QY 1 CCGACACCATCAATGGTGCAAAACCTTTCCGGGTATGCGATGATAGCCCGGAAGAGA 60
D 1 CCGACACCATCAATGGTGCAAAACCTTTCCGGGTATGCGATGATAGCCCGGAAGAGA 60
QY 61 GTCAATTCAAGGTGGTGAATGTGAACACCAAGTATACGATGTGCGAGAGTATGCC 120
D 61 GTCAATTCAAGGTGGTGAATGTGAACACCAAGTATACGATGTGCGAGAGTATGCC 120
QY 121 GTGTCTTTATCAGACCGGTTTCCCGGTGGTGAACACGAGCCAGCCAGCTTCTCGCAAA 180
```

```
D 121 GTGTCTTTATCAGACCGGTTTCCCGGTGGTGAACACGAGCCAGCCAGCTTCTCGCAAA 180
QY 181 CGCGGAAAAAGTGGAAAGCGCGATGCGGAGTGAATTAATTTCCCAACCGCGTGGCAC 240
D 181 CGCGGAAAAAGTGGAAAGCGCGATGCGGAGTGAATTAATTTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAACAGTCTGCTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGC 300
D 241 AACAACTGGCGGGCAACAGTCTGCTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCCGCTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGC 360
D 301 ACGCCGCTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGC 360
QY 361 TGGTGGTGTGATGATAGAACGAGCGCGTGAAGCTGTAAAGCGCGCTGACAAATC 420
D 361 TGGTGGTGTGATGATAGAACGAGCGCGTGAAGCTGTAAAGCGCGCTGACAAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATATCCGCTGGATGACCAAGTGC 480
D 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATATCCGCTGGATGACCAAGTGC 480
QY 481 TTGCTGTGGAAGCTCGCTGCACTAATGTTCCGGGTTAATTTCTTGATGTCTCTGACCA 540
D 481 TTGCTGTGGAAGCTCGCTGCACTAATGTTCCGGGTTAATTTCTTGATGTCTCTGACCA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGTACGCGACTGGCGCTGGAGCATC 600
D 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGTACGCGACTGGCGCTGGAGCATC 600
QY 601 TGGTGCATTTGGGTACCAAGCAAAATCGCGCTTTAGCGGGCCATTAAGTCTGTCTGG 660
D 601 TGGTGCATTTGGGTACCAAGCAAAATCGCGCTTTAGCGGGCCATTAAGTCTGTCTGG 660
QY 661 CGCGTCTGCGCTCGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
D 661 CGCGTCTGCGCTCGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCGACTGGAGTGCCATGTCGGTTTTTCAACAAACCATGCAATGCTGA 780
D 721 CGGAACGGGAAGGCGACTGGAGTGCCATGTCGGTTTTTCAACAAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTCGCATGCTGGTTGCGCAAGTCAAGTGGCGCTGGCGCAA 840
D 781 ATGAGGGCATCGTTCCCACTCGCATGCTGGTTGCGCAAGTCAAGTGGCGCTGGCGCAA 840
QY 841 TGGCGCATTTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGTAGTGGGATAGC 900
D 841 TGGCGCATTTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGTAGTGGGATAGC 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
D 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAAAACCAAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCGAGCGGTGA 1020
D 961 GCCTGCTGGGCAAAACCAAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCGAGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
D 1021 AGGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAATGAGCTGGCAGCAGCAGGTTT 1140
D 1081 CGCAAAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAATGAGCTGGCAGCAGGTTT 1140
QY 1141 CCGACTGGAAGCGGGCAGTGAGCGCAACCGCAATTAATGTAGTTAGTCACTCATTTAG 1200
D 1141 CCGACTGGAAGCGGGCAGTGAGCGCAACCGCAATTAATGTAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAAATTTCTATGTTTTCAGAGCTTATCATGACTGCGAGGTGCAACCAATGCTTCTGGC 1260
```


Db 1201 GCACAAATCTCTAATGTTTGACAGCTTATCATCGACTGCGACGGTGCCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGCAGCCATCGGAAGCTGTGGTATCGCTGTGAGGTGCTAAATCACTGCATATTCG 1320
Db 1261 TCAGGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTAAATCACTGCATATTCG 1320
QY 1321 TGTCTCTCAAGCGGCACCTCCGTTCTGGATTAATGTTTTTGGCGCGACATCATACGGTT 1380
Db 1321 TGTCTCTCAAGCGGCACCTCCGTTCTGGATTAATGTTTTTGGCGCGACATCATACGGTT 1380
QY 1381 CTGGCAAAATATCTCABAATGAGCTGTTGACAAATTAATCATCGCTCGTAAATGTTGGGA 1440
Db 1381 CTGGCAAAATATCTCABAATGAGCTGTTGACAAATTAATCATCGCTCGTAAATGTTGGGA 1440
QY 1441 ATTGTAGCGGATACAAATTTTACACAGAGAAACAGCCAGTCCGTTTAGTGTGTTTTCAAGA 1500
Db 1441 ATTGTAGCGGATACAAATTTTACACAGAGAAACAGCCAGTCCGTTTAGTGTGTTTTCAAGA 1500
QY 1501 GCATTTCACCAACAGGACCAATAGATTATGAAATCTGAAGAGGTAAATCTGTTAATCTGG 1560
Db 1501 GCATTTCACCAACAGGACCAATAGATTATGAAATCTGAAGAGGTAAATCTGTTAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCCGTTAGGTTAAGAAATTCGAGAAAGT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCCGTTAGGTTAAGAAATTCGAGAAAGT 1620
QY 1621 ACCGGAAATTAAGTCAACGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGTT 1680
Db 1621 ACCGGAAATTAAGTCAACGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGTT 1680
QY 1681 GGGCACTGCGGATGGCCCTGACATTAATCTTCTGGGACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GGGCACTGCGGATGGCCCTGACATTAATCTTCTGGGACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGACAAAGGTTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGACAAAGGTTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTACAAACGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTACAAACGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGGGAAACCCGCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGGGAAACCCGCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCTGGATTAAGAACTGAAAGCGAAAGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCTGGATTAAGAACTGAAAGCGAAAGTAAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGACGGGGTTATGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGACGGGGTTATGGTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGATTAAGAGCGTGGCGTGGATACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGATTAAGAGCGTGGCGTGGATACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAAGAAACAAACACATGAATGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAAGAAACAAACACATGAATGAGACACCGATTAC 2160
QY 2161 TCCATCCGAGAGTGGCTTTTAAAGGCGAAACACGCGATGACCATCAACGCGCCGTGG 2220
Db 2161 TCCATCCGAGAGTGGCTTTTAAAGGCGAAACACGCGATGACCATCAACGCGCCGTGG 2220
QY 2221 GCATGGTCCCAACATCCGACACGACGAAAGTGAATGATGGTTAAGCGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCCAACATCCGACACGACGAAAGTGAATGATGGTTAAGCGTACTGCCGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAACGCGCCGAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAACGCGCCGAGT 2340

RESULT 4

US-09-765-555-17

; Sequence 17, Application US/09765555

; Publication No. US20030037355A1

; GENERAL INFORMATION:

; APPLICANT: The Scripps Research Institute

; TITLE OF INVENTION: Methods and compositions to modulate

QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTTCTCGAAAACCTATCTGTGACTGATGAAGGCTCG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTTCTCGAAAACCTATCTGTGACTGATGAAGGCTCG 2400
QY 2401 GAAGCGGTAAATAAAGACAAAACCGCTGGGTGCGGTAGCGTGAAGTCTTTACGAGGAAAGAG 2460
Db 2401 GAAGCGGTAAATAAAGACAAAACCGCTGGGTGCGGTAGCGTGAAGTCTTTACGAGGAAAGAG 2460
QY 2461 TTGGCGAAAGATTCACGCTATTGCGCGCACCATGGAATAAACCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGCGAAAGATTCACGCTATTGCGCGCACCATGGAATAAACCCAGAAAGGTGAAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGGCTTTCTGCTGATGCGGTGCTAGTGGGTGATCAAGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGGCTTTCTGCTGATGCGGTGCTAGTGGGTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGTCAGACTGTCGATGAAGCCCTCAAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGTCAGACTGTCGATGAAGCCCTCAAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAAATAAATAAACAACCTCGGGATCGAGGAAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAAACAAATAAATAAACAACCTCGGGATCGAGGAAAGGATTCAGAAATTC 2700
QY 2701 GGATCTCTTCTCTGTCGGCCAGGGCGCCCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTCGGCCAGGGCGCCCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTGGTAAAGTCTCTCTCAGAGCTCTCAGCTGGTGGCGCCACACAGCTGATCCAC 2820
Db 2761 CCGGAATGTGGTAAAGTCTCTCTCAGAGCTCTCAGCTGGTGGCGCCACACAGCTGATCCAC 2820
QY 2821 ACAGGTGAAACAAACGCTATAATGCCAGAGTGGGCAAAATCTTTTACGCAAGTCCAGCAAC 2880
Db 2821 ACAGGTGAAACAAACGCTATAATGCCAGAGTGGGCAAAATCTTTTACGCAAGTCCAGCAAC 2880
QY 2881 CTGCTGGCCATCAACGCTACTATCTGGCGAAGCCATACAAATGTCCAGAAATGTTGGC 2940
Db 2881 CTGCTGGCCATCAACGCTACTATCTGGCGAAGCCATACAAATGTCCAGAAATGTTGGC 2940
QY 2941 AAGTCTTCTCTCGGTCTGACAAATCTCGTCCGCGACCAACAGTACTCACACCGGGGAGAAG 3000
Db 2941 AAGTCTTCTCGGTCTGACAAATCTCGTCCGCGACCAACAGTACTCACACCGGGGAGAAG 3000
QY 3001 CCCTATGCTGTCGCGAAATGTGTAAGTCTTTCAGCGCGCAGGATAACTGTTGGTGGCCAC 3060
Db 3001 CCCTATGCTGTCGCGAAATGTGTAAGTCTTTCAGCGCGCAGGATAACTGTTGGTGGCCAC 3060
QY 3061 CAGCGTACCCACACGCGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACGCGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
QY 3121 CAGCGCGGCAACTGTCGCGAGCCATCAACGCTACTGCGGAGAGCCATACAAATGT 3180
Db 3121 CAGCGCGGCAACTGTCGCGAGCCATCAACGCTACTGCGGAGAGCCATACAAATGT 3180
QY 3181 CCAGAATGTGGCAAGTCTTCTCGTCTGCAATCTCGTCCGCGACCAACGCTACTCAC 3240
Db 3181 CCAGAATGTGGCAAGTCTTCTCGTCTGCAATCTCGTCCGCGACCAACGCTACTCAC 3240
QY 3241 ACCGGTAAAAAAACCTAGTGGCCAGCGCGCCAGTACCCCGTACGACGTTCCGGACTACGCT 3300
Db 3241 ACCGGTAAAAAAACCTAGTGGCCAGCGCGCCAGTACCCCGTACGACGTTCCGGACTACGCT 3300

; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m4 and zinc finger
; OTHER INFORMATION: protein ZFPm4
US-09-765-555-17

Query Match 97.4%; Score 3215.2; DB 10; Length 3300;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 3247; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 1 CGACACCATCGAATGCTGCAAAACCTTTCCGGGTATGCGATAGAGCGCCGGGAAGA 60
DB 1 CGACACCATCGAATGCTGCAAAACCTTTCCGGGTATGCGATAGAGCGCCGGGAAGA 60

QY 61 GTCAATTGAGGGTGTGAATGTGAACACAGTAACGTTATACGATGTCGACAGTATGCCG 120
DB 61 GTCAATTGAGGGTGTGAATGTGAACACAGTAACGTTATACGATGTCGACAGTATGCCG 120

QY 121 GTGTCTTATCAGACCGTTTCCCGGTGTGAACACGAGCCAGCCAGCTTTCTCGAAAA 180
DB 121 GTGTCTTATCAGACCGTTTCCCGGTGTGAACACGAGCCAGCCAGCTTTCTCGAAAA 180

QY 181 CGCGGAAAGTGAAGCGGCGATGGCGGAGCTCAATTACATCCCAACCGCTGGCAC 240
DB 181 CGCGGAAAGTGAAGCGGCGATGGCGGAGCTCAATTACATCCCAACCGCTGGCAC 240

QY 241 AACAACTGGCGGCAACAGTCTGTTGCTGATTTGGGGTTGCCACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGCAACAGTCTGTTGCTGATTTGGGGTTGCCACCTCCAGTCTGGCCCTGC 300

QY 301 AGCGCGGTGCAAAATGTCGCGGCGATTAATATCTCGCGCGCATCAATCTGGTGGTCCAGCG 360
DB 301 AGCGCGGTGCAAAATGTCGCGGCGATTAATATCTCGCGCGCATCAATCTGGTGGTCCAGCG 360

QY 361 TGGTGTGTCGATGTAGAACGAGCGGCTGCAAGCGCTGTAAAGCGGCGTGCACAATC 420
DB 361 TGGTGTGTCGATGTAGAACGAGCGGCTGCAAGCGCTGTAAAGCGGCGTGCACAATC 420

QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTCCGCTGGATGACCAAGGATGCCA 480
DB 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTCCGCTGGATGACCAAGGATGCCA 480

QY 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTATTTCTTGATGTCCTGACACAGA 540
DB 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTATTTCTTGATGTCCTGACACAGA 540

QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGCATGCGCGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGCATGCGCGTGGAGCATC 600

QY 601 TGGTGTGCTGTCACAGCAAAATCGCGCTGTATAGCGGCGCATTTAAGTTCTGTCTCGG 660
DB 601 TGGTGTGCTGTCACAGCAAAATCGCGCTGTATAGCGGCGCATTTAAGTTCTGTCTCGG 660

QY 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720

QY 721 CGGAACGGGAAGCGACTGGAGTGCATGTCGGTTTTCACAAACCATGCAAAATGCTGA 780
DB 721 CGGAACGGGAAGCGACTGGAGTGCATGTCGGTTTTCACAAACCATGCAAAATGCTGA 780

DB 721 CGGAACGGGAAGCGACTGGAGTGCATGTCGGTTTTCACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGCGCTGGGCGCAA 840
DB 781 ATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGCGCTGGGCGCAA 840
QY 841 TGGCGGCATTACCGAGTCCGGGCTGGCGTGGTGGCGGATATCTCGTAGTGGGATACG 900
DB 841 TGGCGGCATTACCGAGTCCGGGCTGGCGTGGTGGCGGATATCTCGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTAAACCAACCATCAAAACGAGATTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTAAACCAACCATCAAAACGAGATTTC 960
QY 961 GCCTGCTGGGGCAAAACACAGCGTGCACCGCTTGTCTGCAACTCTCTCAGGCGCCAGGCGTGA 1020
DB 961 GCCTGCTGGGGCAAAACACAGCGTGCACCGCTTGTCTGCAACTCTCTCAGGCGCCAGGCGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCTCTCATCTGGTGAAGAAAGAAACCAACCTGSGGCCAATA 1080
DB 1021 AGGGCAATCAGCTGTTGCCCTCTCATCTGGTGAAGAAAGAAACCAACCTGSGGCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCAATTAATGAGCTGGCAGACGAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCAATTAATGAGCTGGCAGACGAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
DB 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAATCTCATGTTTGACAGCTTATCATCGACTGCAAGGTGCACCAATGCTTCTGGCG 1260
DB 1201 GCACAATCTCATGTTTGACAGCTTATCATCGACTGCAAGGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGCAGCCATCGGAAGCTGTGGTATGCGGTGTCAGGTCTGTAATCACTGCATAATTCG 1320
DB 1261 TCAGCAGCCATCGGAAGCTGTGGTATGCGGTGTCAGGTCTGTAATCACTGCATAATTCG 1320
QY 1321 TGTGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCAACGCTT 1380
DB 1321 TGTGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCAACGCTT 1380
QY 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
DB 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
QY 1441 ATTGTAGCGGATAACAAATTTACACAGGAACAGCCAGTCCGTTTAGTGTTCACGA 1500
DB 1441 ATTGTAGCGGATAACAAATTTACACAGGAACAGCCAGTCCGTTTAGTGTTCACGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAGGTAAACTGGTAACTCTGG 1560
DB 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAGGTAAACTGGTAACTCTGG 1560
QY 1561 ATTTAAGCGGATAAAGGCTATAACGCTCTCGTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTTAAGCGGATAAAGGCTATAACGCTCTCGTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATGGAAGAGAAATTCACAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATGGAAGAGAAATTCACAGGTT 1680
QY 1681 CGGGCAACTGGCGATGGCCCTGACATATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
DB 1681 CGGGCAACTGGCGATGGCCCTGACATATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCCGGAAGAGGTTCCAGGACAAGCTGTAT 1800
DB 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCCGGAAGAGGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCGGTACGTTACAAACCGGAGCTGATTTGCTTACCCGATCGCTGT 1860
DB 1801 CCGTTTACCTGGATGCGGTACGTTTACAAACCGGAGCTGATTTGCTTACCCGATCGCTGT 1860

```
QY 1861 GAAGCGTTATCGCTGATTTTATACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTTATACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTTCACCTGGCGCTGATTGCTGCTGACGGGGGTATTGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGGCGCTGATTGCTGCTGACGGGGGTATTGCGTTCAAG 2040
QY 2041 TATGAAAAACGGCAAGTACGATTAAGAACTGAAGCGGTGGATTAACCGTGGCGCGAAACGG 2100
Db 2041 TATGAAAAACGGCAAGTACGATTAAGAACTGAAGCGGTGGATTAACCGTGGCGCGAAACGG 2100
QY 2101 GGTCTGACCTTCCCTGGTTGACTGATTAAGAACTGAAGCGGTGGATTAACCGTGGCGCGAAACGG 2160
Db 2101 GGTCTGACCTTCCCTGGTTGACTGATTAAGAACTGAAGCGGTGGATTAACCGTGGCGCGAAACGG 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCAAATCGACACCGAGCAAGTGAATTATGGTGTAAACGGTACTCGCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACCGAGCAAGTGAATTATGGTGTAAACGGTACTCGCGACCTTC 2280
QY 2281 AAGGTCAACCATCCAAACCGTGGTGGCTGCTGAGCGCAGGTATTAACGGCCCGCAGT 2340
Db 2281 AAGGTCAACCATCCAAACCGTGGTGGCTGCTGAGCGCAGGTATTAACGGCCCGCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAACTATCTGCTGACTGATGAAGGCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAACTATCTGCTGACTGATGAAGGCTG 2400
QY 2401 GAAGCGTTAATAAGCAAAACCGCTGGGTGGCTGAGCTGAAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGTTAATAAGCAAAACCGCTGGGTGGCTGAGCTGAAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCAGCTATTGCGCCACCATGGAAGCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGCTATTGCGCCACCATGGAAGCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTGATGCCGTGCTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTGATGCCGTGCTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGATGATGATGAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGATGATGATGAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACCAACCAACAATAACAATAACAACCTCCGGATCGAGGCGAAGATTTTCAGAAATC 2700
Db 2641 AACCAACCAACAATAACAATAACAACCTCCGGATCGAGGCGAAGATTTTCAGAAATC 2700
QY 2701 GGATCTCTTCTCTGTGGCCAGCGGCCCTCGAGCGCCGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTGGCCAGCGGCCCTCGAGCGCCGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTGTAGTCTTCTCTCAGAGCTCTCACTGTGTGCGCCACCGAGCTACCCAC 2820
Db 2761 CCGGAATGTGTAGTCTTCTCTCAGAGCTCTCACTGTGTGCGCCACCGAGCTACCCAC 2820
QY 2821 ACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGTGCGCCATCAACCGACTCATCTGGCGAGAGGCCATACAAATGTCCAGATGTGGC 2940
Db 2881 CTGTGCGCCATCAACCGACTCATCTGGCGAGAGGCCATACAAATGTCCAGATGTGGC 2940
```

```
QY 2941 AAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTACTACACCGGGGAGAAG 3000
Db 2941 AAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTACTACACCGGGGAGAAG 3000
QY 3001 CCTATGCTTGTCCGGAATGTGTAAGTCTTTCAGCGCGCAGCATAACCTGGTGGCCAC 3060
Db 3001 CCTATGCTTGTCCGGAATGTGTAAGTCTTTCAGCGCGCAGCATAACCTGGTGGCCAC 3060
QY 3061 CAGCGTACCACACGCGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAATCTTTTACG 3120
Db 3061 CAGCGTACCACACGCGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAATCTTTTACG 3120
QY 3121 CAGCGCGCCACCTGGCGCAGCATCAACGCACTCATACTGGCGAGAGCCATACAAATGT 3180
Db 3121 CAGCGCGCATCAACCTGGCGCAGCATCAACGCACTCATACTGGCGAGAGCCATACAAATGT 3180
QY 3181 CCAGAAATGTGGCAAGTCTTTCTCTCGGTCTGCAAAATCTCGTCCGGCACCAACGTACTCAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTTCTCAACTTCAGGCCATTTGGTCCGTCAACCAACGTACTCAC 3240
QY 3241 ACCGTTAAAAAACTAGTGGCGCGCGCGCATACCGGTACCGACTTCGGGACTACGCT 3300
Db 3241 ACCGTTAAAAAACTAGTGGCGCGCGCGCATACCGGTACCGACTTCGGGACTACGCT 3300
```

RESULT 5

```
US-09-765-555-16
; Sequence 16, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m3 and zinc finger
; OTHER INFORMATION: protein ZFPm3
US-09-765-555-16
```

```
Query Match 97.2%; Score 3207.2; DB 10; Length 3300;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTCAGGTGGTGAATGTGAAAACCGTAACGTTATACGATGTCGAGAGTATGCCG 120
Db 61 GTCAATTCAGGTGGTGAATGTGAAAACCGTAACGTTATACGATGTCGAGAGTATGCCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCCGGTGTGAACCCAGCCAGCCAGCTTCTTCCGCAAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCCGGTGTGAACCCAGCCAGCCAGCTTCTTCCGCAAAA 180
QY 181 CGCGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
QY 241 ACAACCTGGCGGCAACACAGTGGTGTGCTGATTTGGCGTTGCCACCTCAGTCTGGCCCTGC 300
```

Db 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACCTCCAGTCGCGCCCTGC 300
Qy 301 ACGCGCGTCGCAAAATFTGCGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGTCGCAAAATFTGCGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
Qy 361 TGGTGTGTGATGTTAGAACGAAGCGGCTGGAAGCCTGTAAAGCGGCGTGCACAAATC 420
Db 361 TGGTGTGTGATGTTAGAACGAAGCGGCTGGAAGCCTGTAAAGCGGCGTGCACAAATC 420
Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTATCCGTGGATGACCAAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTATCCGTGGATGACCAAGGATGCCA 480
Qy 481 TTGCTGTGAAAGCTGCTGCACTAATGTTCCGGGCTTATTTCTGATGTCCTGACAGCA 540
Db 481 TTGCTGTGAAAGCTGCTGCACTAATGTTCCGGGCTTATTTCTGATGTCCTGACAGCA 540
Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACGCACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACGCACTGGCGGTGGAGCATC 600
Qy 601 TGGTCGCTTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCATTTAAGTTCCTGCTCGG 660
Db 601 TGGTCGCTTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCATTTAAGTTCCTGCTCGG 660
Qy 661 CGCGTCTGCGTCTGCGTGGCTGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGCGTGGCTGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAACGCACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAACGCACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAAACGATCAGATGGCGCTGGCGCAA 840
Db 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAAACGATCAGATGGCGCTGGCGCAA 840
Qy 841 TCGCGCCCATACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGTGGGATAG 900
Db 841 TCGCGCCCATACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGTGGGATAG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAAAAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAAAAGGATTTTC 960
Qy 961 GCCTGCTGGGGCAAAACCGAGTGGACCGCTTGTGCAACTCTCTCAGGSCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGGCAAAACCGAGTGGACCGCTTGTGCAACTCTCTCAGGSCCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGATTCATTAATGCACTGGCAACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGATTCATTAATGCACTGGCAACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCACTGAGCGCAACGCAATTAATGCTGAGTTAGCTCACTATTAG 1200
Db 1141 CCCGACTGGAAGCGGCACTGAGCGCAACGCAATTAATGCTGAGTTAGCTCACTATTAG 1200
Qy 1201 GCACAATTCATGTTTGAAGCTTATCATGCACTGCAGCGGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTCATGTTTGAAGCTTATCATGCACTGCAGCGGTGCACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGCAAGTTCGTAATCACTGATCAATTCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGCAAGTTCGTAATCACTGATCAATTCG 1320
Qy 1321 TGTGCTGAAGGGCACTCCCGTTCTGGATATGTTTTTGGCCGACATCAACGGTT 1380
Db 1321 TGTGCTGAAGGGCACTCCCGTTCTGGATATGTTTTTGGCCGACATCAACGGTT 1380

Qy 1381 CTGGCAAAATATCTGAAATAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTGTGGA 1440
Db 1381 CTGGCAAAATATCTGAAATAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTGTGGA 1440
Qy 1441 ATTGTAGCGGATAAACAATTTTACACAGAAACAGCAGTCCGTTTACGTGTTTTCACGA 1500
Db 1441 ATTGTAGCGGATAAACAATTTTACACAGAAACAGCAGTCCGTTTACGTGTTTTCACGA 1500
Qy 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAGGTAAACTGGTAACTCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAGGTAAACTGGTAACTCTGG 1560
Qy 1561 ATTAAACGGCGATTAAGGCTTAAACCGTCTCGTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAAACGGCGATTAAGGCTTAAACCGTCTCGTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTTAAAGTCACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTTCCACAGGTT 1680
Db 1621 ACCGGAATTTAAAGTCACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTTCCACAGGTT 1680
Qy 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTTGGGCAACGACCCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTTGGGCAACGACCCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGGCTAGTTTACAAACGCAAGCTGATTTGCTTACCCGATCCCTGTT 1860
Db 1801 CCGTTTACCTGGGATGGCTAGTTTACAAACGCAAGCTGATTTGCTTACCCGATCCCTGTT 1860
Qy 1861 GAAAGCTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAAGCTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATTTCAAC 1980
Qy 1981 CTGCAAGAACGCTACTTCACTGGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACGCTACTTCACTGGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAACGCGCAAGTACGACATTAAGACGCTGGCGCTGGATAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGCGCAAGTACGACATTAAGACGCTGGCGCTGGATAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGTTGACCTGATTAATAACAAACATGAAATGCAAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTGACCTGATTAATAACAAACATGAAATGCAAGACACCGATTAC 2160
Qy 2161 TCCATCGCAAGTGGCTTTTAAAGCGGAAACAGCGATGACCATCAACGGCCGCTGG 2220
Db 2161 TCCATCGCAAGTGGCTTTTAAAGCGGAAACAGCGATGACCATCAACGGCCGCTGG 2220
Qy 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAATGTTGTAACGCTCTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAATGTTGTAACGCTCTGCGGACCTTC 2280
Qy 2281 AAGGTCACCATCAAAACCGTTTCTGGCGTGGTGAAGAGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTCACCATCAAAACCGTTTCTGGCGTGGTGAAGAGTATTAACGCCGCCAGT 2340
Qy 2341 CCGAACAAGAGCTGGCAAAAGAGTTTCTCGAAAATCTATCTGCTGACTGATGAAGGCTCG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTTCTCGAAAATCTATCTGCTGACTGATGAAGGCTCG 2400
Qy 2401 GAAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAGCTTCTTACGAGGAAG 2460
Db 2401 GAAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAGCTTCTTACGAGGAAG 2460

QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGGACTGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGGACTGGCGTGGAGCATC 600
QY 601 TGGTCGCATTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db 601 TGGTCGCATTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGCTGCGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTGCGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTTTCAACAACCATGCAATGCTGA 780
Db 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTTTCAACAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCTCACTGCGATGCTGTGGTTCACACGATCGCTGGCGCGAA 840
Db 781 ATGAGGGCATCGTTCCTCACTGCGATGCTGTGGTTCACACGATCGCTGGCGCGAA 840
QY 841 TGCGCGCCATTACCGAGTCCGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db 841 TGCGCGCCATTACCGAGTCCGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACACCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACACCATCAACAGGATTTTC 960
QY 961 GCCTCTGGGGCAACACGAGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Db 961 GCCTCTGGGGCAACACGAGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGGCCCGTCTCACTGGTGAAGAAAAACCCCTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCCGTCTCACTGGTGAAGAAAAACCCCTGGCGGCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGCGAGCTGGCAGCAGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGCGAGCTGGCAGCAGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCGGTGCAACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCGGTGCAACCAATGCTTCTGGCG 1260
QY 1261 TCAGGCAGCCATCGGAAGCTGTGGTATGCGCTGCGAGGTGTAATCATCTGCAATATTCG 1320
Db 1261 TCAGGCAGCCATCGGAAGCTGTGGTATGCGCTGCGAGGTGTAATCATCTGCAATATTCG 1320
QY 1321 TGTCCCTCAAGCGCACTCCCGTTCTGGATATGTTTTTGGCGCGACATCAACGGTT 1380
Db 1321 TGTCCCTCAAGCGCACTCCCGTTCTGGATATGTTTTTGGCGCGACATCAACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
QY 1441 ATTGTAGCGGATAACAATTTTCAACAGAAACAGCCAGTCCGTTTAGTGTTCAGGA 1500
Db 1441 ATTGTAGCGGATAACAATTTTCAACAGAAACAGCCAGTCCGTTTAGTGTTCAGGA 1500
QY 1501 GCACCTTCCCAACAGGACCATAGATTATGAAAACTGAAGAGTAAACTGGTATCTGG 1560
Db 1501 GCACCTTCCCAACAGGACCATAGATTATGAAAACTGAAGAGTAAACTGGTATCTGG 1560
QY 1561 ATTTAACGGCGATAAAGGCTTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTTAACGGCGATAAAGGCTTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAAGAGAAATTTCCACAGGTT 1680

Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTTCCACAGGTT 1680
QY 1681 CGCGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTGTGGTGGCTAC 1740
Db 1681 CGCGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTGTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCCTGTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCCTGTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGTACGTTACAAACGCAAGCTGATTGTTACCCGATGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGTACGTTACAAACGCAAGCTGATTGTTACCCGATGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGTAAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTACCTGCCGCTGATTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTGCCGCTGATTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATCAAAAACGCAAGTACGACATTAAAGACGCTGGCGTGGATTAACGCTGGCGGCAAGCG 2100
Db 2041 TATGAAAACGCAAGTACGACATTAAAGACGCTGGCGTGGATTAACGCTGGCGGCAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGTTGACCTGATTAAAAAACAACATGAATGACAGACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGTTGACCTGATTAAAAAACAACATGAATGACAGACCGGATTAC 2160
QY 2161 TCCATCCGAGAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCCGAGAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGTTCCAAATCGACACCAACGCAAGTGAATTAATGTTGTAACGCTACTGCGACCTTC 2280
Db 2221 GCATGTTCCAAATCGACACCAACGCAAGTGAATTAATGTTGTAACGCTACTGCGACCTTC 2280
QY 2281 AAGGTCACACCATCAAAACCGTTCTGTTGGCTGTGAGCGCAGGTATTAACCGCCAGT 2340
Db 2281 AAGGTCACACCATCAAAACCGTTCTGTTGGCTGTGAGCGCAGGTATTAACCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACCTATCTGCTGACTGATGAGGTCG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACCTATCTGCTGACTGATGAGGTCG 2400
QY 2401 GAAGCGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGCGGAAAGATCAACGTTATTCGCCCAACCATGGAACACGCCAGAAAAGGTGAAATCATG 2520
Db 2461 TTGCGGAAAGATCAACGTTATTCGCCCAACCATGGAACACGCCAGAAAAGGTGAAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGCTGCTACTGCGGTCATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGCTGCTACTGCGGTCATCAACGCC 2580
QY 2581 GCCAGCGTCTGACAGTCTGATGAAGCCCTGAAGACGCGACACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTCTGATGAAGCCCTGAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAATAACAATAACAACCAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAAACAATAACAATAACAACCAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
QY 2701 GGATCCTCTTCTGTGGGCCACGCGGCCCTCGAGGCCGG 2741

Db 2701 GGATCCTCCAGTCCCGGCAAGCCGCCCTGGTGGGAGG 2741

RESULT 7

```

US-10-149-472-5
; Sequence 5, Application US/10149472
; Publication No. US20040029204A1
; GENERAL INFORMATION:
; APPLICANT: GOUBIN-GRAMATICA, FRANCOISE
; APPLICANT: DUCOMMUN, BERNARD
; APPLICANT: PREVOST, GREGOIRE
; TITLE OF INVENTION: METHOD FOR OBTAINING HUMAN CDC25 PHOSPHATASES AND METHOD
; TITLE OF INVENTION: FOR IDENTIFYING HUMAN CDC25 PHOSPHATASE MODULATORS
; FILE REFERENCE: bml-427.065
; CURRENT APPLICATION NUMBER: US/10/149,472
; CURRENT FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: PCT/FR00/03496
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: FR 99/15722
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: FR 00/06683
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: FR 00/12008
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 8101
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: JM109 / pMAL - Hs Cdc25C strain
US-10-149-472-5

```

Query Match	82.0%	Score 2705.8	DB 17	Length 8101
Best Local Similarity	99.9%	Pred. No. 0		
Matches 2707	Conservative 0	Mismatches 2	Indels 0	Gaps 0
QY	1	CCGACACCATCGAATGTCGAAACACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA	60	
Db	1	CCGACACCATCGAATGTCGAAACACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA	60	
QY	61	GTCAATTTCAGGCTGGTGAATGTGAAACCCAGTAACTGATATCGATGTGCGAGAGTATGCCG	120	
Db	61	GTCAATTTCAGGCTGGTGAATGTGAAACCCAGTAACTGATATCGATGTGCGAGAGTATGCCG	120	
QY	121	GTGCTCTTATCAGACCGTTTCCGCGTGTGTAACCGAGCCGAGCCACCGTTTCTCGGAAAA	180	
Db	121	GTGCTCTTATCAGACCGTTTCCGCGTGTGTAACCGAGCCGAGCCACCGTTTCTCGGAAAA	180	
QY	181	CGCGGAAAAAGTGGAGCGCGGATGCGCGAGCTGAAATTACATTCCCAACCGCGTGGCAC	240	
Db	181	CGCGGAAAAAGTGGAGCGCGGATGCGCGAGCTGAAATTACATTCCCAACCGCGTGGCAC	240	
QY	241	AACAACTGGCGGGCAACACGTGCTGCTGATTTGGCGTTGGCACTCTCAGTCTTGGCCCTGC	300	
Db	241	AACAACTGGCGGGCAACACGTGCTGCTGATTTGGCGTTGGCACTCTCAGTCTTGGCCCTGC	300	
QY	301	ACGCGCGTTCGCAAAATTGTCGCGGCAATTAATCTCTCGCGCGATCAACTCGGTGCCACGC	360	
Db	301	ACGCGCGTTCGCAAAATTGTCGCGGCAATTAATCTCTCGCGCGATCAACTCGGTGCCACGC	360	
QY	361	TGGTGTGTGATGTTAGACGAAGCGGCGTCGAAGCCTGTAAAGCGCGGTGCACAATC	420	
Db	361	TGGTGTGTGATGTTAGACGAAGCGGCGTCGAAGCCTGTAAAGCGCGGTGCACAATC	420	
QY	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAACATATCCGCTGGATGACACGAGTGC	480	
Db	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAACATATCCGCTGGATGACACGAGTGC	480	
QY	481	TTGCTGTGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGATCTTGACCGA	540	

Db	481	TTGCTGTGGAAAGCTGCCTGCACTAAATGTTCCGGCGGTTATTCTTCTGAATGTTCTCTGACCCAGA	541
Qy	541	CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600
Db	541	CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600
Qy	601	TGGTCGCAATTCGGGTACACAGCAAAATCGCGCTGTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG	660
Db	601	TGGTCGCAATTCGGGTACACAGCAAAATCGCGCTGTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG	660
Qy	661	CGCGTCTCGGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Db	661	CGCGTCTCGGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Qy	721	CGGAAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA	780
Db	721	CGGAAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA	780
Qy	781	ATGAGGGCATGTTCCCACTCGGATGCTGGTTGCCAAGATCAGATGGCGCTGGGCGCAA	840
Db	781	ATGAGGGCATGTTCCCACTCGGATGCTGGTTGCCAAGATCAGATGGCGCTGGGCGCAA	840
Qy	841	TGCGCGCATTAACACGAGTCCGGGCTGGCGGTTGGTGGCGATATCTCGTACTGGGATACG	900
Db	841	TGCGCGCATTAACACGAGTCCGGGCTGGCGGTTGGTGGCGATATCTCGTACTGGGATACG	900
Qy	901	ACGATACCGAAGACAGTCTATGTTATATCCGCCGCTTTAAACACATCAACAGAGATTTTC	960
Db	901	ACGATACCGAAGACAGTCTATGTTATATCCGCCGCTTTAAACACATCAACAGAGATTTTC	960
Qy	961	GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTTGA	1020
Db	961	GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTTGA	1020
Qy	1021	AGGCAATCAGCTGTTGCCGCTCTCACTGTGTAAGTAAAGAAAACCAACCTGGCGGCCAATA	1080
Db	1021	AGGCAATCAGCTGTTGCCGCTCTCACTGTGTAAGTAAAGAAAACCAACCTGGCGGCCAATA	1080
Qy	1081	CGCAAAACCGCTCTCCCGCGGCTTGCCCGATTCATTAAATGCAGCTGGCAGCAGGTTT	1140
Db	1081	CGCAAAACCGCTCTCCCGCGGCTTGCCCGATTCATTAAATGCAGCTGGCAGCAGGTTT	1140
Qy	1141	CCGACTGGAAAGCGGCGAGTGACGCCAACGCCAATTAATGTAGTTAGCTCACTCATTAG	1200
Db	1141	CCGACTGGAAAGCGGCGAGTGACGCCAACGCCAATTAATGTAGTTAGCTCACTCATTAG	1200
Qy	1201	GCACAATTCATGTTTGACAGCTTATCATCGCTGACCGGTGGCCAAATGCTCTCGGCTGGG	1260
Db	1201	GCACAATTCATGTTTGACAGCTTATCATCGCTGACCGGTGGCCAAATGCTCTCGGCTGGG	1260
Qy	1261	TCAGGACGCATCGGAAGCTGTGGTATGGCTGTGCGAGTCTGTAATCACTGCAATAATTCG	1320
Db	1261	TCAGGACGCATCGGAAGCTGTGGTATGGCTGTGCGAGTCTGTAATCACTGCAATAATTCG	1320
Qy	1321	TGTCGCTCAAGGCGCATCCCGTCTCGATAATGTTTTTGGCCGACATCAATACCGTT	1380
Db	1321	TGTCGCTCAAGGCGCATCCCGTCTCGATAATGTTTTTGGCCGACATCAATACCGTT	1380
Qy	1381	CTGCAAAATTTCTGAAATGAGCTGTTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA	1440
Db	1381	CTGCAAAATTTCTGAAATGAGCTGTTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA	1440
Qy	1441	ATTGTGAGCGGATAACAAATTTACACAGGAAAACAGCCAGTCCGTTTATAGTTGTTTTACGA	1500
Db	1441	ATTGTGAGCGGATAACAAATTTACACAGGAAAACAGCCAGTCCGTTTATAGTTGTTTTACGA	1500
Qy	1501	GCATTTCAACCAACAGGACCATAGATTATGAAAACCTGAAGAGGTAAATCGTAAATCTCGG	1560
Db	1501	GCATTTCAACCAACAGGACCATAGATTATGAAAACCTGAAGAGGTAAATCGTAAATCTCGG	1560
Qy	1561	ATTAAACGGGATAAAGGCTATAACGGTCTCGCTGAAATCGGTGAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAAACGGGATAAAGGCTATAACGGTCTCGCTGAAATCGGTGAAGAAATTCGAGAAAGAT	1620

QY 1621 ACCGGAATTAAAGTCAACCGTTGAGCATCCGGATAAACTGGAAAGAGAAATTCCTCCACAGGTT 1680
Db |||||||
QY 1621 ACCGGAATTAAAGTCAACCGTTGAGCATCCGGATAAACTGGAAAGAGAAATTCCTCCACAGGTT 1680
Db |||||||
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACGCGCTTTGGTGGCTAC 1740
Db |||||||
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACGCGCTTTGGTGGCTAC 1740
Db |||||||
QY 1741 GCTCAATCTGGCCTGTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db |||||||
QY 1801 CCGTTTACCTGGATGCGGTAGCTTAAACGCAAGCTGATTTGCTTACCGGATCGCTGTT 1860
Db |||||||
QY 1801 CCGTTTACCTGGATGCGGTAGCTTAAACGCAAGCTGATTTGCTTACCGGATCGCTGTT 1860
Db |||||||
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGTCGCCGAACCCGCCAAACCTGGGAA 1920
Db |||||||
QY 1921 GAGATCCCGGCTCGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db |||||||
QY 1981 CTGCAAGAACCGTACTTCACTGGCGGTGATTTGCTGACGCGGGTATGCGTTCAAG 2040
Db |||||||
QY 1981 CTGCAAGAACCGTACTTCACTGGCGGTGATTTGCTGACGCGGGTATGCGTTCAAG 2040
Db |||||||
QY 2041 TATGAAACCGCAAGTACGACATTAAGACGTGGCGGTGGATTAACGCTGGCGGAAAGG 2100
Db |||||||
QY 2041 TATGAAACCGCAAGTACGACATTAAGACGTGGCGGTGGATTAACGCTGGCGGAAAGG 2100
Db |||||||
QY 2101 GGTCTGACCTTCTGCTGATTAATAAAGCAAAACATCAATCAATCAATCAATCAATCAAT 2160
Db |||||||
QY 2161 TCCATCGAAGAGTGCCTTTAATAAGCGGAAACAGCGATCAACATCAACCGCCCGTGG 2220
Db |||||||
QY 2161 TCCATCGAAGAGTGCCTTTAATAAGCGGAAACAGCGATCAACATCAACCGCCCGTGG 2220
Db |||||||
QY 2221 GCATGGTCCACATCGACACGCAAGTGAATTAAGTGAACGGTACTGCGGACCTTC 2280
Db |||||||
QY 2221 GCATGGTCCACATCGACACGCAAGTGAATTAAGTGAACGGTACTGCGGACCTTC 2280
Db |||||||
QY 2281 AAGGTCACACATCAACACCGTTCTGTCGCTGAGCGAGGTATTAACCGCCCGCAGT 2340
Db |||||||
QY 2281 AAGGTCACACATCAACACCGTTCTGTCGCTGAGCGAGGTATTAACCGCCCGCAGT 2340
Db |||||||
QY 2341 CCGAACAAAGAGTGGCAAAAGAGTTCTCGAAACTATCTGACTGATGAAGGTCTG 2400
Db |||||||
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGTCCGCTAGCGTGAAGTCTTACGAGGAAG 2460
Db |||||||
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGTCCGCTAGCGTGAAGTCTTACGAGGAAG 2460
Db |||||||
QY 2461 TTGGGAAAGATCCACGTTATTCGCCCAACATGGAAGAACCCAGAAAGGTGAATCATG 2520
Db |||||||
QY 2461 TTGGGAAAGATCCACGTTATTCGCCCAACATGGAAGAACCCAGAAAGGTGAATCATG 2520
Db |||||||
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGTGGTACTGCGGTGATCAACGCC 2580
Db |||||||
QY 2581 GCCAGGGTGTGATGATGATGAAGCCCTGAAAGAGCGCGAGACTAATTCAGAGTCTG 2640
Db |||||||
QY 2581 GCCAGGGTGTGATGATGATGAAGCCCTGAAAGAGCGCGAGACTAATTCAGAGTCTG 2640
Db |||||||
QY 2641 AACAAACAAACAATAAACAATAACACCTCGGATCGAGGAGGATTTTCAGAAATTC 2700
Db |||||||
QY 2641 AACAAACAAACAATAAACAATAACACCTCGGATCGAGGAGGATTTTCAGAAATTC 2700
Db |||||||

QY 2701 GGATCCTCT 2709
Db |||||||
QY 2701 GGATCCTCT 2709
Db |||||||
RESULT 8
US-10-343-859-8
; Sequence 8, Application US/10343859
; Publication No. US20040110161A1
; GENERAL INFORMATION:
; APPLICANT: Nanogen Recognition GMBH
; TITLE OF INVENTION: Method for detecting mutations in
; FILE REFERENCE: 612,406-033
; CURRENT APPLICATION NUMBER: US/10/343,859
; CURRENT FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: PCT/EP01/08127
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 10038237.1
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 6648
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: plasmid
; OTHER INFORMATION: pMAL-c2x
US-10-343-859-8

Query Match 81.8%; Score 2701; DB 19; Length 6648;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2704; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGGAAAACCTTTCCGGGTATGGCATGATAGCCCGGAAGAGA 60
Db |||||||
QY 1 CCGACACCATCGAATGGTGGAAAACCTTTCCGGGTATGGCATGATAGCCCGGAAGAGA 60
Db |||||||
QY 61 GTCATTCAGGGTGGTGAATGTGAAACCAAGTAAAGTATAGATGTCGACAGTATGCG 120
Db |||||||
QY 61 GTCATTCAGGGTGGTGAATGTGAAACCAAGTAAAGTATAGATGTCGACAGTATGCG 120
Db |||||||
QY 121 GTGTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCACGTTTCTGCGAAAA 180
Db |||||||
QY 121 GTGTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCACGTTTCTGCGAAAA 180
Db |||||||
QY 181 CCGGGAAAAAGTGGAAAGCGGCGATGGCGGAGCTGAATTAATTCCTCCACCGCGTGGCAC 240
Db |||||||
QY 181 CCGGGAAAAAGTGGAAAGCGGCGATGGCGGAGCTGAATTAATTCCTCCACCGCGTGGCAC 240
Db |||||||
QY 241 AACAACTCGCGGCAAAACAGTCTGTTGTTGGGTTGGCCACCTCCAGTCTGGCCCTGC 300
Db |||||||
QY 241 AACAACTCGCGGCAAAACAGTCTGTTGTTGGGTTGGCCACCTCCAGTCTGGCCCTGC 300
Db |||||||
QY 301 ACGCGCGTTCGCAAAATTTCTCGCGCGATTAATCTCGCGCGGATCAATTCCTGGGTGCGAGCG 360
Db |||||||
QY 301 ACGCGCGTTCGCAAAATTTCTCGCGCGATTAATCTCGCGCGGATCAATTCCTGGGTGCGAGCG 360
Db |||||||
QY 361 TGGTGGTTCGATGGTAGAAAGAGCGGCTGGAAGCCCTGTAAGCGGCGGTGCAATTC 420
Db |||||||
QY 361 TGGTGGTTCGATGGTAGAAAGAGCGGCTGGAAGCCCTGTAAGCGGCGGTGCAATTC 420
Db |||||||
QY 421 TTCTCGCGCAACCGCTCAGTGGCTGATTAATTAATCTCGCGCGGATCAATTCCTGGATGCAAGGATGCA 480
Db |||||||
QY 421 TTCTCGCGCAACCGCTCAGTGGCTGATTAATTAATCTCGCGCGGATCAATTCCTGGATGCAAGGATGCA 480
Db |||||||
QY 481 TTGCTGTGGAAGTCTGCTGCACTAATGTTCCGGCGTATTTCTTGATGTTCTGACCCAGA 540
Db |||||||
QY 481 TTGCTGTGGAAGTCTGCTGCACTAATGTTCCGGCGTATTTCTTGATGTTCTGACCCAGA 540
Db |||||||
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGAGCGGTACCGGACTGGGCGTGGAGCATC 600
Db |||||||

Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGAGCATC 600
Qy 601 TGGTCGCAATTGGGTACACAGCAAAATCGCGCTGTAGCGGCCCATTAAGCTTCTGCTCGG 660
Db 601 TGGTCGCAATTGGGTACACAGCAAAATCGCGCTGTAGCGGCCCATTAAGCTTCTGCTCGG 660
Qy 661 CGGCTCTGGCTCTGGCTGGCTGGCTAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db 661 CGGCTCTGGCTCTGGCTGGCTGGCTAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Qy 721 CGGAACGGGAAGCGGACGTGAGTGCATGTCGGGTTTCAAAACCAATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGCGGACGTGAGTGCATGTCGGGTTTCAAAACCAATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTCGCAACGATCAGATGGCGCTGGCGCAA 840
Db 781 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTCGCAACGATCAGATGGCGCTGGCGCAA 840
Qy 841 TGGCGGCCATTACCGAGTCCGGCTGGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGGCCATTACCGAGTCCGGCTGGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTATATCCCGCGTTAAACACCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTATATCCCGCGTTAAACACCATCAACAGGATTTTC 960
Qy 961 GCCTGCTGGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGTGA 1020
Db 961 GCCTGCTGGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGGCGGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCGGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGGCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTAATTAATGAGCTGACGACGAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTAATTAATGAGCTGACGACGAGGTTT 1140
Qy 1141 CCCGCTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGAGTGAAGTCACTCATTTAG 1200
Db 1141 CCCGCTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGAGTGAAGTCACTCATTTAG 1200
Qy 1201 GCACAAATCTCATGTTGACAGCTTATCATGACGTGACGCTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTGACAGCTTATCATGACGTGACGCTGACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGACAGTGTGTAATCACTGCATAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGACAGTGTGTAATCACTGCATAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGGTTCTGGATATGTTTTTGGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGGTTCTGGATATGTTTTTGGCGCGACATCAATACGGTT 1380
Qy 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
Qy 1441 ATTGTGAGCGGATTAACAAATTTCAACAGGAACAGCCAGTCCGTTTAGTGTTTTTACGA 1500
Db 1441 ATTGTGAGCGGATTAACAAATTTCAACAGGAACAGCCAGTCCGTTTAGTGTTTTTACGA 1500
Qy 1501 GCATTTCAACCAAGGACCATAGATTATGAACCTGAAGGTAACCTGGTAATCTGG 1560
Db 1501 GCATTTCAACCAAGGACCATAGATTATGAACCTGAAGGTAACCTGGTAATCTGG 1560
Qy 1561 ATTAACGGCGATTAAGGCTATAACGGTCTCGCTGAAGTCCGTAAGAAATTCGAGAAGAT 1620
Db 1561 ATTAACGGCGATTAAGGCTATAACGGTCTCGCTGAAGTCCGTAAGAAATTCGAGAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCACAGGTT 1680

Qy 1681 GGGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACGACCGCTTTGGTGGCTAC 1740
Db 1681 GGGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCCTTTGGCTGAAATCACCCCGGACAAAGGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCCTTTGGCTGAAATCACCCCGGACAAAGGTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACTTGGGATGCGGTACGTTTACAACGGCAAGCTGATTTGCTTACCGGATCGCTGTT 1860
Db 1801 CCGTTTACTTGGGATGCGGTACGTTTACAACGGCAAGCTGATTTGCTTACCGGATCGCTGTT 1860
Qy 1861 GAACGGTTATCGCTGATTTATACAAAGATCTGCTGCGAACCCGCCCAAAAACCTGGGAA 1920
Db 1861 GAACGGTTATCGCTGATTTATACAAAGATCTGCTGCGAACCCGCCCAAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGTAAAGCGCGCTGTATTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGTAAAGCGCGCTGTATTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTTCACCTGGCCGCTGATTTGCTGCTGACCGGGGTTATGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGGCCGCTGATTTGCTGCTGACCGGGGTTATGGTTCAAG 2040
Qy 2041 TATGAAAAACGGCAAGTACGACATTTAAAGACGTGGCGGTGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAAACGGCAAGTACGACATTTAAAGACGTGGCGGTGATTAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAACTGCTTTTAAAGGCGAAACAGCGATGACCATCAACCGCCCGTGG 2220
Db 2161 TCCATCGCAGAACTGCTTTTAAAGGCGAAACAGCGATGACCATCAACCGCCCGTGG 2220
Qy 2221 GCATGCTCAAACATCGACACAGCAAGTGAATTTATGGTGTAAACGCTGCTGCGACCTTC 2280
Db 2221 GCATGCTCAAACATCGACACAGCAAGTGAATTTATGGTGTAAACGCTGCTGCGACCTTC 2280
Qy 2281 AAGGCTCAACCATTCAAAACCGTTTGGCTGCTGAGCGCAGGTATTAACCGCCGAGT 2340
Db 2281 AAGGCTCAACCATTCAAAACCGTTTGGCTGCTGAGCGCAGGTATTAACCGCCGAGT 2340
Qy 2341 CCGAAACAAAGAGCTGCAAAAGAGTTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGCAAAAGAGTTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGTTTAAATAAGACAAAACCGCTGGGTGCGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
Db 2401 GAAGCGTTTAAATAAGACAAAACCGCTGGGTGCGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
Qy 2461 TTTGGCAAAAGATCCAGTATTTGCGGCCACATGGAAGAAACGCCGAGAAAGGTGAATCATG 2520
Db 2461 TTTGGCAAAAGATCCAGTATTTGCGGCCACATGGAAGAAACGCCGAGAAAGGTGAATCATG 2520
Qy 2521 CCGAAACATCCGAGATGTCGCTTTCTGATGCGTGGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAAACATCCGAGATGTCGCTTTCTGATGCGTGGTACTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGAGCGCGCAGACTAATTTGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGAGCGCGCAGACTAATTTGAGCTCG 2640
Qy 2641 AACACAAACAATAACAAATTAACACACTCTCGGATCGAGGAGGATTTTCAGAAATTC 2700
Db 2641 AACACAAACAATAACAAATTAACACACTCTCGGATCGAGGAGGATTTTCAGAAATTC 2700
Qy 2701 GGATCTCT 2709
Db 2701 GGATCTCT 2709

RESULT 9

US-10-343-859-9
; Sequence 9, Application US/10343859
; Publication No. US20040110161A1
; GENERAL INFORMATION:
; APPLICANT: Nanogen Recognomics GMBH
; TITLE OF INVENTION: Method for detecting mutations in
; FILE OF INVENTION: nucleotide sequences
; FILE REFERENCE: 612,406-033
; CURRENT APPLICATION NUMBER: US/10/343,859
; PRIOR FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: PCT/EP01/08127
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 10038237.1
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 9191
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: pMALc2x-muts
; OTHER INFORMATION: plasmid
US-10-343-859-9

Query Match 81.8%; Score 2698; DB 19; Length 9191;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2701; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY	1	CCGACACCATCGAATGGTGCAGAAACCTTTCGGGGTATGGCATGATAGGCCCGGAGAGA	60
DB	1	CCGACACCATCGAATGGTGCAGAAACCTTTCGGGGTATGGCATGATAGGCCCGGAGAGA	60
QY	61	GTCATTCAGGTGGTGAATGTGAACACAGTAACTATACCATGTCGACAGATGCGG	120
DB	61	GTCATTCAGGTGGTGAATGTGAACACAGTAACTATACCATGTCGACAGATGCGG	120
QY	121	GTGCTCTTATCAGACCGGTTTCGGCGTGTGAACACAGCCAGCCACGTTTCTGGGAAA	180
DB	121	GTGCTCTTATCAGACCGGTTTCGGCGTGTGAACACAGCCAGCCACGTTTCTGGGAAA	180
QY	181	CCGGGAAAAGTGAAGCGGATGGCGAGTGAATTAATCCCAACCGCGTGGCAC	240
DB	181	CCGGGAAAAGTGAAGCGGATGGCGAGTGAATTAATCCCAACCGCGTGGCAC	240
QY	241	ACAACCTGCGGGCAACAGTGTGCTGATTTGGGCTTCCACCTCCAGTCTGGCCCTGC	300
DB	241	ACAACCTGCGGGCAACAGTGTGCTGATTTGGGCTTCCACCTCCAGTCTGGCCCTGC	300
QY	301	ACGCGCGTGCAGAAATGTTCGGCGGATTAATCTCGCCCGATCAACTGGGTGCCAGG	360
DB	301	ACGCGCGTGCAGAAATGTTCGGCGGATTAATCTCGCCCGATCAACTGGGTGCCAGG	360
QY	361	TGGTGTGTCGATGTGATAGAACGCGGTGTAAGCGGCTGTAAGCGGCTGCAATC	420
DB	361	TGGTGTGTCGATGTGATAGAACGCGGTGTAAGCGGCTGTAAGCGGCTGCAATC	420
QY	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGGCTGATGACAGGATGCCA	480
DB	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGGCTGATGACAGGATGCCA	480
QY	481	TTGCTGTGAAGTGTGCTCAGTAAATTTTCGGCGTTATTTCTGATGTCCTGACCGA	540
DB	481	TTGCTGTGAAGTGTGCTCAGTAAATTTTCGGCGTTATTTCTGATGTCCTGACCGA	540
QY	541	CACCCATCAACAGTATTTTCTCCATGAAGACGGTACGCGTGGCGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTTTCTCCATGAAGACGGTACGCGTGGCGTGGAGCATC	600
QY	601	TGGTGCATTTGGTTCACAGCAAAATCGCGCTGTATAGCGGGCCCAATTAAGTTCTGCTCGG	660
DB			

DB	601	TGGTGCATTTGGTTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG	660
QY	661	CGCGTCTCGCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	661	CGCGTCTCGCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA	780
DB	721	CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA	780
QY	781	ATGAGGGCATCGTTCCTCCACTGCGATGCTGGTTCGCAACGATCAGATGCGCTGGCGCAA	840
DB	781	ATGAGGGCATCGTTCCTCCACTGCGATGCTGGTTCGCAACGATCAGATGCGCTGGCGCAA	840
QY	841	TGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG	900
DB	841	TGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCATCAACAGGATTTTC	960
DB	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCATCAACAGGATTTTC	960
QY	961	GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTCTCAACTCTCTCAGGCGCAGGCGTGA	1020
DB	961	GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTCTCAACTCTCTCAGGCGCAGGCGTGA	1020
QY	1021	AGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAAAACACCCCTGGCGCCCAATA	1080
DB	1021	AGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAAAACACCCCTGGCGCCCAATA	1080
QY	1081	CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGTGGCAGCAGAGTTT	1140
DB	1081	CGCAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGTGGCAGCAGAGTTT	1140
QY	1141	CCCGACTGGAAGCGGCGAGTGGCGCAACGCAATTAATGTAGTTAGTCTCACTCATTAG	1200
DB	1141	CCCGACTGGAAGCGGCGAGTGGCGCAACGCAATTAATGTAGTTAGTCTCACTCATTAG	1200
QY	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCGGTGCAACCAATGTTCTGGCG	1260
DB	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCGGTGCAACCAATGTTCTGGCG	1260
QY	1261	TCAGCAGCGCATCGAAGCTGTGTATGGCTGTGAGGTGCTGAATCACTGATCAATTCG	1320
DB	1261	TCAGCAGCGCATCGAAGCTGTGTATGGCTGTGAGGTGCTGAATCACTGATCAATTCG	1320
QY	1321	TGTCGCTCAAGCGCACCTCCGTTCTGGATTAATGTTTGGCGCGACATCAACGGTT	1380
DB	1321	TGTCGCTCAAGCGCACCTCCGTTCTGGATTAATGTTTGGCGCGACATCAACGGTT	1380
QY	1381	CTGGCAAAATATTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGGA	1440
DB	1381	CTGGCAAAATATTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGGA	1440
QY	1441	ATTGTGAGCGGATAACAATTTTCAACAGAAACAGCCAGTCCGTTAGGTGTTTTCACGA	1500
DB	1441	ATTGTGAGCGGATAACAATTTTCAACAGAAACAGCCAGTCCGTTAGGTGTTTTCACGA	1500
QY	1501	GCACCTTCAACCAAGGACCATAGATTATGAACTGGAAGAGTAACTGGTGAATCTGG	1560
DB	1501	GCACCTTCAACCAAGGACCATAGATTATGAACTGGAAGAGTAACTGGTGAATCTGG	1560
QY	1561	ATTAACGGCGATAAAGGCTATAACCGCTCTCGCTGAAGTCCGTAAGAAATTCAGAAAGAT	1620
DB	1561	ATTAACGGCGATAAAGGCTATAACCGCTCTCGCTGAAGTCCGTAAGAAATTCAGAAAGAT	1620
QY	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCGGATAAATGGAAGAGAAATTCACAGGTT	1680
DB	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCGGATAAATGGAAGAGAAATTCACAGGTT	1680
QY	1681	GGGGCAATGGCGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC	1740
DB	1681	GGGGCAATGGCGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC	1740

```
QY 1741 GCTCAATCTGGCTGTGTGGCTGAATTCACCCCGGACAAAGCGTTCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGTGGCTGAATTCACCCCGGACAAAGCGTTCAGGACAAGCTGTAT 1800
QY 1801 CGTTTACCTGGATGCGGTAGCTTTACACGGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
Db 1801 CGTTTACCTGGATGCGGTAGCTTTACACGGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
QY 1861 GAAAGCGTTATCGCTGATTTATAACAAGATCTGTCGCGAACCCGCCAAACCTCGGAA 1920
Db 1861 GAAAGCGTTATCGCTGATTTATAACAAGATCTGTCGCGAACCCGCCAAACCTCGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATAAGAACTGAAGCGAAAGGTAAAGAGCGCGTGTATTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAGAACTGAAGCGAAAGGTAAAGAGCGCGTGTATTTCAAC 1980
QY 1981 CTGCAAGACCGTACTTACCTGGCGGTGATGCTGTGACGGGGTGTATGCGTTCAAG 2040
Db 1981 CTGCAAGACCGTACTTACCTGGCGGTGATGCTGTGACGGGGTGTATGCGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGTTGAGCGTGGATACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGTTGAGCGTGGATACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAAGAACTGAAGCAATGAATGACAGACCGGATPAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAAGAACTGAAGCAATGAATGACAGACCGGATPAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAAGCGGCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAAGCGGCCGTGG 2220
QY 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTAAGTGTAAAGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTAAGTGTAAAGTACTGCCGACCTTC 2280
QY 2281 AAGGTCAACCATCAACCGTTTGTGGCGTGTGAGCGGAGGATTAACCGCGCCAGT 2340
Db 2281 AAGGTCAACCATCAACCGTTTGTGGCGTGTGAGCGGAGGATTAACCGCGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAACATATCTGCTGACTGATGAGGTTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAACATATCTGCTGACTGATGAGGTTCTG 2400
QY 2401 GAAAGCGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAAGCGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGGGAAGATCCACGTTATTTGCCGCCACCATGGAACCGCCGAGAAAGGTGAATCATG 2520
Db 2461 TTGGGGAAGATCCACGTTATTTGCCGCCACCATGGAACCGCCGAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCGTGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCGTGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGAGATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGAGATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACCAACAAACAAATACAAATAACAACTCGGATCGAGGGAAGGATTTTCAAGTTTC 2700
Db 2641 AACCAACAAACAAATACAAATAACAACTCGGATCGAGGGAAGGATTTTCAAGTTTC 2700
QY 2701 GGATCC 2706
Db 2701 GGATCC 2706
```

RESULT 10

US-10-263-153-40

; Sequence 40, Application US/10263153

```
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 7112
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3189)
; OTHER INFORMATION: pMBP-c2X-Toxop30del11 (52-214aa)
; US-10-263-153-40
```

```
Query Match 81.5%; Score 2690.4; DB 18; Length 7112;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGTGTGCAAAACCTTTCGCGGTATGGCATATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGTGTGCAAAACCTTTCGCGGTATGGCATATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTCAGGCTGGTCAATGTGAAACCAAGTAACTATACGATGTGCGAGATGATGCCG 120
Db 61 GTCAATTCAGGCTGGTCAATGTGAAACCAAGTAACTATACGATGTGCGAGATGATGCCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCGCGTGTGTAACCAAGCGCAGCCAGCTTCTTCGCAAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCGCGTGTGTAACCAAGCGCAGCCAGCTTCTTCGCAAAA 180
QY 181 CGCGGNAAGTGAAGCGCGCATGCGGAGCTGAAATTAATTCCCAACCGCGTGGCAC 240
Db 181 CGCGGNAAGTGAAGCGCGCATGCGGAGCTGAAATTAATTCCCAACCGCGTGGCAC 240
QY 241 AACCACTGGCGGCAAAACAGTCTGCTGATTGCGCGTGGCACCTCCAGTCTGCGCCTGC 300
Db 241 AACCACTGGCGGCAAAACAGTCTGCTGATTGCGCGTGGCACCTCCAGTCTGCGCCTGC 300
QY 301 ACGGCGCGTCGCAAAATGTGCGCGGATTAATAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db 301 ACGGCGCGTCGCAAAATGTGCGCGGATTAATAATCTCGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGTGTGTGTCGATGTTAGAACGAGCGCGTGAAGCGCTGAAGCGGCTGAAGCGGCTGCAATC 420
Db 361 TGTGTGTGTCGATGTTAGAACGAGCGCGTGAAGCGCTGAAGCGGCTGCAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGCTGATCATTAATCTATCGCTGGATGACAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGCTGATCATTAATCTATCGCTGGATGACAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAATTTGTTCCGCGGTTATTTCTTGATGTCTTGACAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAATTTGTTCCGCGGTTATTTCTTGATGTCTTGACAGA 540
QY 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 600
QY 601 TGTGTCGATTTGGTCAACAGCAAAATCGCGTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Db 601 TGTGTCGATTTGGTCAACAGCAAAATCGCGTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
QY 661 CGGCTCTGCGTCTGGCTGGCGATTAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
```

Db 661 |||||CGCGTCTGCGTCTGCGTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAAACGGGAAAGCGACATGGAGTGCATGTCGGGTTTTCAACAAACCAATGCAAAATGCTGA 780
Db 721 CGGAAACGGGAAAGCGACATGGAGTGCATGTCGGGTTTTCAACAAACCAATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATGTTCCCACTGGCATGCTGGTTGCCAACGATCAGATGGCGTGGCGCAA 840
Db 781 ATGAGGGCATGTTCCCACTGGCATGCTGGTTGCCAACGATCAGATGGCGTGGCGCAA 840
Qy 841 TGCGGCGCATTAACAGATCCGGGCTCGCGTGGTGGCGATATCTCGTGGTGGGATACG 900
Db 841 TGCGGCGCATTAACAGATCCGGGCTCGCGTGGTGGCGATATCTCGTGGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAACCACCACTCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAACCACCACTCAAAACAGGATTTTC 960
Qy 961 GCCTCTGGGGCAAAACGAGCTGGAACCGTCTGCTGCAATCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTCTGGGGCAAAACGAGCTGGAACCGTCTGCTGCAATCTCTCAGGGCCAGGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAAGAAACCAACCTGGGCGCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAAGAAACCAACCTGGGCGCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGTGGCGACGACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGTGGCGACGACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGAGTGGTACCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGAGTGGTACCTCACTCATTTAG 1200
Qy 1201 GCACAATTCATGTTTCACAGCTTATCATCGACTGCGAGGTGCACCAATGCTTCTGGG 1260
Db 1201 GCACAATTCATGTTTCACAGCTTATCATCGACTGCGAGGTGCACCAATGCTTCTGGG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAATCACTGCATAAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAATCACTGCATAAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTCTCGATATGTTTTTTCGCGCAATCAATCAACGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTCGATATGTTTTTTCGCGCAATCAATCAACGTT 1380
Qy 1381 CTGGCAATATCTCAAGTACGCTTTCACATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAATATCTCAAGTACGCTTTCACATTAATCATCGGCTCGTATATGTTGGA 1440
Qy 1441 ATTGTAGCGGATAACAATTTTCACAGAAACAGCCAGTCCGTTTAGTGTGTTTCACGA 1500
Db 1441 ATTGTAGCGGATAACAATTTTCACAGAAACAGCCAGTCCGTTTAGTGTGTTTCACGA 1500
Qy 1501 GCACCTTCAACAAAGGACCATAGATTAATGAACCTGAAGAGGTAACCTGGTAATCTGG 1560
Db 1501 GCACCTTCAACAAAGGACCATAGCATATGAATAATCGAAGAGGTAACCTGGTAATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATACCGTCTCGTGAAGTGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATACCGTCTCGTGAAGTGGTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACGTTGAGCATCCGATTAACCTGGAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGTTGAGCATCCGATTAACCTGGAGAGAAATTCACAGGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCGGGCACACAGCCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCGGGCACACAGCCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTTGGCTGAATTCACCCCGGCAAGAGCGTTCCAGGACAAAGCTGTAT 1800
|||||

Db 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTTACTCGGATGCGGTACGTTTCAACCGGCAAGCTGATTGTTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTTACTCGGATGCGGTACGTTTCAACCGGCAAGCTGATTGTTTACCCGATCGCTGTT 1860
Qy 1861 GAAGGTTATCGCTGATTATTAACAAGATCTGTCGCGAAACCCGCCAAACCTGGGAA 1920
Db 1861 GAAGGTTATCGCTGATTATTAACAAGATCTGTCGCGAAACCCGCCAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCGCTGGATAAAGAACTGAACGAAAGTAAAGAGCGGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAACGAAAGTAAAGAGCGGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCTGATCTTCACTGCGCTGATGTTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCTGATCTTCACTGCGCTGATGTTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAGACGCAAGTACGACATTAAGACGTTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAGACGCAAGTACGACATTAAGACGTTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGCTGCTGATTAATAAACAACATGAAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGCTGATTAATAAACAACATGAAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCCCTTAAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCCCTTAAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGCTCAACATCGACACAGCAAAAGTAAATTAATGCTGTAACCGTACTGCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACAGCAAAAGTAAATTAATGCTGTAACCGTACTGCGGACCTTC 2280
Qy 2281 AAGGTCACACCATCAACCGTCTGTTGGCGTGTGAGCGAGGTATTAACCGCCCGAGT 2340
Db 2281 AAGGTCACACCATCAACCGTCTGTTGGCGTGTGAGCGAGGTATTAACCGCCCGAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACATATCTGCTGACTGATGAAGCTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACATATCTGCTGACTGATGAAGCTCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGTAGCGTGAAGTCTTACGAGGAAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGTAGCGTGAAGTCTTACGAGGAAAG 2460
Qy 2461 TTGCGAAAGATCCACGTTATTCGCGCCACTATGGAACACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGCGAAAGATCCACGTTATTCGCGCCACTATGGAACACGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCGCTTTCTGGTATGCGGCTGCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCGCTTTCTGGTATGCGGCTGCTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Qy 2641 AACACACACAAATTAACAATAACCAACCTCGGGATCGAGGAAAGGATTCAGAAATTC 2700
Db 2641 AACACACACAAATTAACAATAACCAACCTCGGGATCGAGGAAAGGATTCAGAAATTC 2700

RESULT 11
US-10-263-153-30
; Sequence 30, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.

1861 GAAGGTTATCGCTGATTTATACAAAGATCTGTCGCCAAACCCGCCAAAAACCTGGGAA 1920
1921 GAGATCCCGCGCTGGATAAGAACTGAAGAGCAAGAGTAAGAGCGCGCTGATGTTCAAC 1980
1921 GAGATCCCGCGCTGGATAAGAACTGAAGAGCAAGAGTAAGAGCGCGCTGATGTTCAAC 1980
1981 CTGCAAGAACCGTACTTACACCTGGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
1981 CTGCAAGAACCGTACTTACACCTGGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
2041 TATGAAAAACGCAAGTACGACATTAAGACACGTCGGCGTGGATACGCTGGCGCAAGACG 2100
2041 TATGAAAAACGCAAGTACGACATTAAGACACGTCGGCGTGGATACGCTGGCGCAAGACG 2100
2101 GGTCTGACCTTCTCTGTTGACCTGATTAATAAAGGCAACACATGAATGACAGACCGATTAC 2160
2101 GGTCTGACCTTCTCTGTTGACCTGATTAATAAAGGCAACACATGAATGACAGACCGATTAC 2160
2161 TCCATCGAGAAAGTGCCTTTAATAAAGGCAACACATGAATGACAGACCGATTAC 2220
2161 TCCATCGAGAAAGTGCCTTTAATAAAGGCAACACATGAATGACAGACCGATTAC 2220
2221 GCATGTCCTCAACATCGACACGACGCAAGTGAATGATGTTAAACGCTGACCGACCTTC 2280
2221 GCATGTCCTCAACATCGACACGACGCAAGTGAATGATGTTAAACGCTGACCGACCTTC 2280
2281 AAGGTCACACATCCAAACCGTTCGTTGGCGTGTGAGCGAGGATTAAGCGCGCCAGT 2340
2281 AAGGTCACACATCCAAACCGTTCGTTGGCGTGTGAGCGAGGATTAAGCGCGCCAGT 2340
2341 CCGAACAAGAGTGGCAAAAGAGTTCCTCGAATACTATCTGCTGACTGATGAAGTCTG 2400
2341 CCGAACAAGAGTGGCAAAAGAGTTCCTCGAATACTATCTGCTGACTGATGAAGTCTG 2400
2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGCGTAGCGCTGAAGTCTTACGAGGAAG 2460
2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGCGTAGCGCTGAAGTCTTACGAGGAAG 2460
2461 TTGGCGAAAGATCCACGATATTGCCGCCACCATGGAACACGCCAGAAAGGTGAATCATG 2520
2461 TTGGCGAAAGATCCACGATATTGCCGCCACCATGGAACACGCCAGAAAGGTGAATCATG 2520
2521 CCGAACAATCCCGAGATGTCGCTTTCTGGTATGCGCTGCGTACTGCGGTGATCAACGCC 2580
2521 CCGAACAATCCCGAGATGTCGCTTTCTGGTATGCGCTGCGTACTGCGGTGATCAACGCC 2580
2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGGAAGACGCGCAGACTAATTCGAGCTCG 2640
2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGGAAGACGCGCAGACTAATTCGAGCTCG 2640
2641 AACACAAACAATAAACAATAAACAACCTCGGGATCGAGGGAAGNTTTCAGAAATTC 2700
2641 AACACAAACAATAAACAATAAACAACCTCGGGATCGAGGGAAGNTTTCAGAAATTC 2700

RESULT 12
US-10-263-153-35
; Sequence 35, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; FILE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 35
; LENGTH: 7322
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528) ... (3399)
; OTHER INFORMATION: pMBP-c2X-Toxop30del10 (52-284aa)
US-10-263-153-35

Query Match 81.5%; Score 2690.4; DB 18; Length 7322;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 CCGACACATCGAATGTTGCAAAAACCTTTCCCGGTATGGCATGATAGCGCCGGAAGAGA 60
Db 1 CCGACACATCGAATGTTGCAAAAACCTTTCCCGGTATGGCATGATAGCGCCGGAAGAGA 60
Qy 61 GTCAATTCAGGTTGTTGAAACCAAGTAAAGTTATACGATGTCGACAGTATGCG 120
Db 61 GTCAATTCAGGTTGTTGAAACCAAGTAAAGTTATACGATGTCGACAGTATGCG 120
Qy 121 GTGTCCTTATCAGACCGTTTCCCGCGTGTGAAACCAAGCCAGCCAGCCACCGTTTCTGCAAAA 180
Db 121 GTGTCCTTATCAGACCGTTTCCCGCGTGTGAAACCAAGCCAGCCAGCCACCGTTTCTGCAAAA 180
Qy 181 CGCGGAAAAAGTGGAAAGCGCGATGCGGAGTGAATTAATTCCTCCAAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGGAAAGCGCGATGCGGAGTGAATTAATTCCTCCAAACCGCGTGGCAC 240
Qy 241 AACAACTGGCGGGCAACAGTCGTTGATTTGGCGTGGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGCAACAGTCGTTGATTTGGCGTGGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGCGCGTTCGCAAAATGTCGCGCGATTAATCTCGCGCGATCAATCGGCGTGGCACG 360
Db 301 ACGCGCGTTCGCAAAATGTCGCGCGATTAATCTCGCGCGATCAATCGGCGTGGCACG 360
Qy 361 TGGTGGTGTGATGTAAGACGAGCGCGTGAAGCGCTGAAAGCGCGGTGCAACAATC 420
Db 361 TGGTGGTGTGATGTAAGACGAGCGCGTGAAGCGCTGAAAGCGCGGTGCAACAATC 420
Qy 421 TTCTCGCGCAACCGTCACTGAGTGGCTGATTAATTAATATCCGCTGGATGACGAGTGC 480
Db 421 TTCTCGCGCAACCGTCACTGAGTGGCTGATTAATTAATATCCGCTGGATGACGAGTGC 480
Qy 481 TTGCTGTGGAAGTCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACACAGA 540
Db 481 TTGCTGTGGAAGTCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACACAGA 540
Qy 541 CACCATCAACAGATTAATTTTCTCCCATGAAGACGCTACCGACTGGCGGTGGAGCATC 600
Db 541 CACCATCAACAGATTAATTTTCTCCCATGAAGACGCTACCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTGGATTTGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAGTTCGTCTCG 660
Db 601 TGGTGGATTTGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAGTTCGTCTCG 660
Qy 661 CGCGTCTGCGTCTGCGTGGCAAAATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGCGTGGCAAAATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CCGAACCGGAAAGGCACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CCGAACCGGAAAGGCACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGCAATCGTTCCTCCACTCGCATGCTGGTTGCCAAGCATCAGATGCGCTGGCGGCA 840
Db 781 ATGAGGCAATCGTTCCTCCACTCGCATGCTGGTTGCCAAGCATCAGATGCGCTGGCGGCA 840
Qy 841 TGGCGCCATTAACGAGTCCGGCTCGCGTGGTGGCGGATATCTCGTAGTGGGATAG 900
Db 841 TGGCGCCATTAACGAGTCCGGCTCGCGTGGTGGCGGATATCTCGTAGTGGGATAG 900

```
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACCACTCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACCACTCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGGCAAAACCGAGCTGGACCGCTTGCTGCAACTCTCTCAGGCCACAGCGGTGA 1020
Db 961 GCCTGCTGGGGCAAAACCGAGCTGGACCGCTTGCTGCAACTCTCTCAGGCCACAGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTTGCCCGTCTCAGCTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGCCCGTCTCAGCTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGCAGCTGGCAGCAGAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGCAGCTGGCAGCAGAGGTTT 1140
QY 1141 CCCGACTCGGAAGCGGCGAGTGAGCGCAACCGAATTAATGAGTGTAGTCACTCATTTAG 1200
Db 1141 CCCGACTCGGAAGCGGCGAGTGAGCGCAACCGAATTAATGAGTGTAGTCACTCATTTAG 1200
QY 1201 GCACAAATTTCTCATGTTTGACAGCTTATCATCGACTGACAGCTGCACCAATCTTCTGGCG 1260
Db 1201 GCACAAATTTCTCATGTTTGACAGCTTATCATCGACTGACAGCTGCACCAATCTTCTGGCG 1260
QY 1261 TCAGGACGCCATCGGAAGCTGTGATGCTGTGCTGAGGTGCTAAATCACTGCATAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGATGCTGTGCTGAGGTGCTAAATCACTGCATAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACCTCCCGTTCTCGATTAATGTTTTTGGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACCTCCCGTTCTCGATTAATGTTTTTGGCGCGACATCAATACGGTT 1380
QY 1381 CTGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGCTCGTATAATGTGTGGA 1440
Db 1381 CTGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGCTCGTATAATGTGTGGA 1440
QY 1441 ATTGTGAGCGGATTAACAAATTTACAGAGAAACAGCGAGTCCGTTTAGGTGTTTTTCAG 1500
Db 1441 ATTGTGAGCGGATTAACAAATTTACAGAGAAACAGCGAGTCCGTTTAGGTGTTTTTCAG 1500
QY 1501 GCATCTCACCAACAGGACCATGATTAATGAATCTGAAGTGAAGTAACTGTTGATCTGG 1560
Db 1501 GCATCTCACCAACAGGACCATGATTAATGAATCTGAAGTGAAGTAACTGTTGATCTGG 1560
QY 1561 ATTAACCGCGATAAAGCTATAACGGTCTCGTGAAGTTCGTTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACCGCGATAAAGCTATAACGGTCTCGTGAAGTTCGTTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTACCGGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACACAGGTT 1680
Db 1621 ACCGGAATTAAGTACCGGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACACAGGTT 1680
QY 1681 CGGCAACTCGCGATGCGCTGACATTAATCTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTCGCGATGCGCTGACATTAATCTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTTGGCTGAAATCAACCCCGGACAAACGTTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTTGGCTGAAATCAACCCCGGACAAACGTTTCCAGGACAAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTACACGCAAGCTGATTCGTTACCGGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTACACGCAAGCTGATTCGTTACCGGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCAAAAACCTCGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCAAAAACCTCGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAGAGCGCGCTGTGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAGAGCGCGCTGTGTTCAAC 1980
```

```
QY 1981 CTGCAAGAACGCTACTTCACTGCGCCGCTGATTCCTGCTGACGGGGTATTCGCTTCAAG 2040
Db 1981 CTGCAAGAACGCTACTTCACTGCGCCGCTGATTCCTGCTGACGGGGTATTCGCTTCAAG 2040
QY 2041 TATGAAAAACGGCAAGTAGTACGACATTTAAAGACGTTGGGCGTGGATACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAAACGGCAAGTAGTACGACATTTAAAGACGTTGGGCGTGGATACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGTTGATGACCTGATTTAAAAACAAACACATGAATGCAGACACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTGATGACCTGATTTAAAAACAAACACATGAATGCAGACACCGGATTAC 2160
QY 2161 TCCATCCGACAGCTGCTTTTAAAGGCGAACACGCGATGACCATCAACGCGCCGTTGG 2220
Db 2161 TCCATCCGACAGCTGCTTTTAAAGGCGAACACGCGATGACCATCAACGCGCCGTTGG 2220
QY 2221 GCATGTTCCAACTCGACACCAACCGTTTCTGTTGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2280
Db 2221 GCATGTTCCAACTCGACACCAACCGTTTCTGTTGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2280
QY 2281 AAGGTTCAACCATCCAAACCGTTTCTGTTGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGTTCAACCATCCAAACCGTTTCTGTTGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGATGATGAGGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGATGATGAGGTCTG 2400
QY 2401 GAAGCGGTTTAAAGACAAACCGTCTGGTGGCTGAGCGCTGAGCTGAAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTTAAAGACAAACCGTCTGGTGGCTGAGCGCTGAGCTGAAAGTCTTACGAGGAAG 2460
QY 2461 TTGCGGAAAGATCCAGTATTTGCGCCACCATCGAAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGCGGAAAGATCCAGTATTTGCGCCACCATCGAAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCGTTTCTGATATGCGGTGCTGCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCGTTTCTGATATGCGGTGCTGCTGCGGTGATCAACGCC 2580
QY 2581 GCAGCGCTGCTGAGACTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCAGCGCTGCTGAGACTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACCAACAAACAAATTAACAAATTAACAAACCTCGGGATCGAGGAAAGGATTTTCAGAAATTC 2700
Db 2641 AACCAACAAACAAATTAACAAATTAACAAACCTCGGGATCGAGGAAAGGATTTTCAGAAATTC 2700
```

RESULT 13

```
US-10-263-153-25
; Sequence 25, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; FILE REFERENCE: 6984.US.O1
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 7352
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)... (3429)
```

i: OTHER INFORMATION: pMBP-c2X-Toxop30del4C (52-294aa)
US-10-263-153-25

Query Match		81.5%	Score 2690.4	DB 18	Length 7352
Best Local Similarity		99.8%	Pred. No. 0		
Matches 2694		Conservative 0	Mismatches 6	Indels 0	Gaps 0
QY	1	CGGACACCATCGAATGGTGC	AAACCTTTTCGGCGGTATG	GCATGATAGCCCGGAAGAGA	60
DB	1	CGGACACCATCGAATGGTGC	AAACCTTTTCGGCGGTATG	GCATGATAGCCCGGAAGAGA	60
QY	61	GTCAATTCAGGGTGGTGAATG	TGAACACAGTAACTATAC	GATGTCGCAGAGTATCGG	120
DB	61	GTCAATTCAGGGTGGTGAATG	TGAACACAGTAACTATAC	GATGTCGCAGAGTATCGG	120
QY	121	GTGTCCTTATCAGACCGGTT	CCCCGGTGGTGAACAGCC	AGCCGACGCTTTCTCGGAAAA	180
DB	121	GTGTCCTTATCAGACCGGTT	CCCCGGTGGTGAACAGCC	AGCCGACGCTTTCTCGGAAAA	180
QY	181	CGCGGAAAAAGTGAAGCGG	CGATGGCGGAGCTGAATTA	CAATTCGCAACCGGTGGC	240
DB	181	CGCGGAAAAAGTGAAGCGG	CGATGGCGGAGCTGAATTA	CAATTCGCAACCGGTGGC	240
QY	241	AACAACTGCGCGGCAAA	CAGTCTGCTGATTGGCGT	TGCGCACCTCCAGTCTGGCCCTGC	300
DB	241	AACAACTGCGCGGCAAA	CAGTCTGCTGATTGGCGT	TGCGCACCTCCAGTCTGGCCCTGC	300
QY	301	ACGCGCGGTGCGCAAA	TTGTGCGGGGATTAATCT	CGCGCGGATCAACTGGGGTGC	360
DB	301	ACGCGCGGTGCGCAAA	TTGTGCGGGGATTAATCT	CGCGCGGATCAACTGGGGTGC	360
QY	361	TGGTGGTTCGATGTAGAA	CGAAGCGGCTGAAAGCC	CTGTAAGCGGCGGTGCAATC	420
DB	361	TGGTGGTTCGATGTAGAA	CGAAGCGGCTGAAAGCC	CTGTAAGCGGCGGTGCAATC	420
QY	421	TTCTCGCGCAACGCGT	CAGTGGGCTGATCAATTA	CTATCCCGTGGATGACCAAGGATGCCA	480
DB	421	TTCTCGCGCAACGCGT	CAGTGGGCTGATCAATTA	CTATCCCGTGGATGACCAAGGATGCCA	480
QY	481	TTGCTGTGGAACTGCCT	GCACATAATGTTCCGGCGT	TATTTCTGATGCTCTGACCCAGA	540
DB	481	TTGCTGTGGAACTGCCT	GCACATAATGTTCCGGCGT	TATTTCTGATGCTCTGACCCAGA	540
QY	541	CACCCATCAACAGTATTA	TTTCTCCCATGAAGACGG	TACGCGATCGGCGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTA	TTTCTCCCATGAAGACGG	TACGCGATCGGCGTGGAGCATC	600
QY	601	TGGTCGCAATGGGT	CACAGCAAAATCGCGCT	GTTAGCGGGCCCATTAAGTTCTGCTCGG	660
DB	601	TGGTCGCAATGGGT	CACAGCAAAATCGCGCT	GTTAGCGGGCCCATTAAGTTCTGCTCGG	660
QY	661	CGCGTCTGCGTCTGCGT	CGCATAAATCTCACTCG	CAATCAAAATTCAGCCGATAG	720
DB	661	CGCGTCTGCGTCTGCGT	CGCATAAATCTCACTCG	CAATCAAAATTCAGCCGATAG	720
QY	721	CGGAACGGGAAGCGAC	TGGAGTGCATGTCCGGT	TTTCAACAAACCATGCAAAATGCTGA	780
DB	721	CGGAACGGGAAGCGAC	TGGAGTGCATGTCCGGT	TTTCAACAAACCATGCAAAATGCTGA	780
QY	781	ATGAGGGCATCGTTCC	CACTGCGATGGTTGCCA	ACGATCAGATGGCGTGGCGGCA	840
DB	781	ATGAGGGCATCGTTCC	CACTGCGATGGTTGCCA	ACGATCAGATGGCGTGGCGGCA	840
QY	841	TGCGGCGCATTAAC	GAGTCCGGGTGGTGGG	TATCTCGGTAGTGGGATAG	900
DB	841	TGCGGCGCATTAAC	GAGTCCGGGTGGTGGG	TATCTCGGTAGTGGGATAG	900
QY	901	ACGATACCGAAGAC	GATCATGTTATATCCCG	CGGTAAACCAACAGGATTTTC	960
DB	901	ACGATACCGAAGAC	GATCATGTTATATCCCG	CGGTAAACCAACAGGATTTTC	960
QY	961	GCCTGCTGGGGCAAC	CCAGCGTGGACCGCTT	GCTGCAACTCTCTCAGGGCCAGGCGGTGA	1020
DB	961	GCCTGCTGGGGCAAC	CCAGCGTGGACCGCTT	GCTGCAACTCTCTCAGGGCCAGGCGGTGA	1020

DB	961	GCCTGCTGGGGCAAC	CCAGCGTGGACCGCTT	GCTGCAACTCTCTCAGGGCCAGGCGGTGA	1020
QY	1021	AGGGCAATCAGCTGT	TGCGCGCTCACCTG	TGTAAGAAAAAACCAACCTGGCGCCCAATA	1080
DB	1021	AGGGCAATCAGCTGT	TGCGCGCTCACCTG	TGTAAGAAAAAACCAACCTGGCGCCCAATA	1080
QY	1081	CGCAACCGCTCTCC	CGCGGTTGGCGGATTA	TAATGCAAGCTGGCAGCAGAGTTT	1140
DB	1081	CGCAACCGCTCTCC	CGCGGTTGGCGGATTA	TAATGCAAGCTGGCAGCAGAGTTT	1140
QY	1141	CCGACTCGAAAGCGG	CAGTGAGCGCAACG	CAATTAATGTAGTCTCACTCATTTAG	1200
DB	1141	CCGACTCGAAAGCGG	CAGTGAGCGCAACG	CAATTAATGTAGTCTCACTCATTTAG	1200
QY	1201	GCACAATTTCTCAT	GTTCACAGCTTATCAT	CGACTGCGACCAATGCTTCTGGCG	1260
DB	1201	GCACAATTTCTCAT	GTTCACAGCTTATCAT	CGACTGCGACCAATGCTTCTGGCG	1260
QY	1261	TCAGGACGCAATCG	GAAGCTGTGGTATGG	CTGTGAGCTCGTAAATCACTGCATTAATCG	1320
DB	1261	TCAGGACGCAATCG	GAAGCTGTGGTATGG	CTGTGAGCTCGTAAATCACTGCATTAATCG	1320
QY	1321	TGTGCTCAAGCGCA	CTCCCGTTCTGATTA	TGTTTTTGGCGCGACATATAACGGTT	1380
DB	1321	TGTGCTCAAGCGCA	CTCCCGTTCTGATTA	TGTTTTTGGCGCGACATATAACGGTT	1380
QY	1381	CTGGCAAAATTTCT	GAAATGAGCTGTG	TGCAATTAATCATCGGCTCGTATATGTGGA	1440
DB	1381	CTGGCAAAATTTCT	GAAATGAGCTGTG	TGCAATTAATCATCGGCTCGTATATGTGGA	1440
QY	1441	ATTGTGAGCGGATA	CAATTTTCACAGGAA	ACAGCCAGTCCGTTAGTGTGTTTCAAGA	1500
DB	1441	ATTGTGAGCGGATA	CAATTTTCACAGGAA	ACAGCCAGTCCGTTAGTGTGTTTCAAGA	1500
QY	1501	GCACCTTACCACCA	AGGACCATAGATTA	TGAAATCTGAAGAGGTAATCTGTAATCTGG	1560
DB	1501	GCACCTTACCACCA	AGGACCATAGATTA	TGAAATCTGAAGAGGTAATCTGTAATCTGG	1560
QY	1561	ATTAACGCGGATAA	AGGCTATAACGCT	CTCGTGAAGTCGTAAGAAATTCGAGAAAAGAT	1620
DB	1561	ATTAACGCGGATAA	AGGCTATAACGCT	CTCGTGAAGTCGTAAGAAATTCGAGAAAAGAT	1620
QY	1621	ACCGGAATTAAGTC	CACCGTTGAGCAT	CCGGATAAATCTGGAAGAGAAATTCGCCACAGGTT	1680
DB	1621	ACCGGAATTAAGTC	CACCGTTGAGCAT	CCGGATAAATCTGGAAGAGAAATTCGCCACAGGTT	1680
QY	1681	CGGCAACTGCGGAT	TGGCCCTGACATTA	TCTTCTGGGCACACGCGTTTGGTGGCTAC	1740
DB	1681	CGGCAACTGCGGAT	TGGCCCTGACATTA	TCTTCTGGGCACACGCGTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCT	TGCTGAAATCAAC	CCCGGACAAAGCGTTCAGGACAAGCTGTAT	1800
DB	1741	GCTCAATCTGGCT	TGCTGAAATCAAC	CCCGGACAAAGCGTTCAGGACAAGCTGTAT	1800
QY	1801	CGGTTTACCTGGAT	TGCGGTACAA	CGGCAAGCTGATTTGTTACCCGATCCGCTGT	1860
DB	1801	CGGTTTACCTGGAT	TGCGGTACAA	CGGCAAGCTGATTTGTTACCCGATCCGCTGT	1860
QY	1861	GAAAGGTTATCGCT	GATTTATAA	CAAGATCTGCTCCGAAACCCGCAAAACCTGGGAA	1920
DB	1861	GAAAGGTTATCGCT	GATTTATAA	CAAGATCTGCTCCGAAACCCGCAAAACCTGGGAA	1920
QY	1921	GAGATCCCGGCT	TGATATAAGAA	ACTGAAAGCGAAGCTGAAGCGCGCTGATGTTCAAC	1980
DB	1921	GAGATCCCGGCT	TGATATAAGAA	ACTGAAAGCGAAGCTGAAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTA	CTTCACTTGGCC	GCTGATGCTGACCGGGGTTATGCGTTCAAG	2040
DB	1981	CTGCAAGAACCGTA	CTTCACTTGGCC	GCTGATGCTGACCGGGGTTATGCGTTCAAG	2040
QY	2041	TATGAAAACGCGCA	AGTACGATTA	AAAGACGTTGGCGTGGATTAACGCTGGCGGAAAGCG	2100
DB	2041	TATGAAAACGCGCA	AGTACGATTA	AAAGACGTTGGCGTGGATTAACGCTGGCGGAAAGCG	2100

```
QY 2101 GGTCTGACCTTCTCTGGTTGACCTGATTTAAACAAACACATGAATGACACACGATTAAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTTAAACAAACACATGAATGACACACGATTAAC 2160
QY 2161 TCCATCGAGAAAGCTGCGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGGTGG 2220
Db 2161 TCCATCGAGAAAGCTGCGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGGTGG 2220
QY 2221 GCATGGTCCACATCGACACAGCAAGCAAGTGAATTAATGTTGTAACCGTACTGCGACCTTC 2280
Db 2221 GCATGGTCCACATCGACACAGCAAGCAAGTGAATTAATGTTGTAACCGTACTGCGACCTTC 2280
QY 2281 AAGGGTCAACCATCAACACCGTTTGGTGGCGTGTGAGCGAGGTATTAACCGCCCAAGT 2340
Db 2281 AAGGGTCAACCATCAACACCGTTTGGTGGCGTGTGAGCGAGGTATTAACCGCCCAAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACATATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACATATCTGCTGACTGATGAAGGTCTG 2400
QY 2401 GAAGCGGTAAATAAGACAAACCGCTGGGTGCGGTAGCGTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTAAATAAGACAAACCGCTGGGTGCGGTAGCGTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGATTTGCGGCCACCATGGAACAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTTGCGGCCACCATGGAACAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGAGAGTCCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGAGAGTCCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACAGTCTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTCTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACACAAACAATTAACAATAACAACAACTCGGGATCGAGGAAGATTTTCAGAAATTC 2700
Db 2641 AACACAAACAATTAACAATAACAACAACTCGGGATCGAGGAAGATTTTCAGAAATTC 2700

RESULT 14
US-10-263-153-20
; Sequence 20, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN.
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 7370
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3447)
; OTHER INFORMATION: pMBP-c2X-Toxo30del13C (52-300aa)
US-10-263-153-20
```

Query Match 81.58; Score 2690.4; DB 18; Length 7370;
Best Local Similarity 99.84; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
QY 1 CCGACACCATCGAATGATGCTGCAAAACCTTTCCGCGGTATGGCATGATAGCGCCCGAAGAGA 60
Db 1 CCGACACCATCGAATGATGCTGCAAAACCTTTCCGCGGTATGGCATGATAGCGCCCGAAGAGA 60
QY 61 GTCAATTCAGGGTGGTGAATGTGAACACAGTTAAGCTTATACGATGTCGACAGAGTATCCG 120
Db 61 GTCAATTCAGGGTGGTGAATGTGAACACAGTTAAGCTTATACGATGTCGACAGAGTATCCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCGCGTGTGTGAACACAGGCGCAGCCAGTTTCTCGGAAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCGCGTGTGTGAACACAGGCGCAGCCAGTTTCTCGGAAAA 180
QY 181 CCGGGGAAAGTGAAGCGCGCATGCGCGAGCTGAATTACATTCACAAACCGGTGGCAC 240
Db 181 CCGGGGAAAGTGAAGCGCGCATGCGCGAGCTGAATTACATTCACAAACCGGTGGCAC 240
QY 241 AACAACTGGCGGGCAAAACAGTCTGCTGATTGGCGTTGCGACCTCCAGTCTTGCCCTTCG 300
Db 241 AACAACTGGCGGGCAAAACAGTCTGCTGATTGGCGTTGCGACCTCCAGTCTTGCCCTTCG 300
QY 301 ACGGCCCGTTCGCAAAATTTGTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCACG 360
Db 301 ACGGCCCGTTCGCAAAATTTGTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCACG 360
QY 361 TGGTGGTGTGATGAGTGAAGCGCGCTGGAAGCTGTGAAGCGCGGTGCAACATC 420
Db 361 TGGTGGTGTGATGAGTGAAGCGCGCTGGAAGCTGTGAAGCGCGGTGCAACATC 420
QY 421 TTCTCGCGCAACCGTCAAGTGGCTGATCAATTAACCTATCCGCTGGATGACACAGATGCCA 480
Db 421 TTCTCGCGCAACCGTCAAGTGGCTGATCAATTAACCTATCCGCTGGATGACACAGATGCCA 480
QY 481 TTCTCGTGGAGCTGCTGCACTAAATTTCCGCGGTATTTCTTGATGTCTTGACACAGA 540
Db 481 TTCTCGTGGAGCTGCTGCACTAAATTTCCGCGGTATTTCTTGATGTCTTGACACAGA 540
QY 541 CACCCATCAACAGTATTTTCTCCCATGAAGACGCTGACGACCTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTTCTCCCATGAAGACGCTGACGACCTGGGCGTGGAGCATC 600
QY 601 TGGTGGTGTGATGAGTGAAGTGGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db 601 TGGTGGTGTGATGAGTGAAGTGGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CCGGTCTGCGTCTGGCTGTCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CCGGTCTGCGTCTGGCTGTCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CCGAACGGGAAGCGACTGGAGTGCCATGTCGGGTTTTCACAAACCATGCAAAATGCTGA 780
Db 721 CCGAACGGGAAGCGACTGGAGTGCCATGTCGGGTTTTCACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCATCTGGATGTCGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
Db 781 ATGAGGGCATCGTTCCCATCTGGATGTCGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
QY 841 TGGCGCCATTACCGAGTCCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 900
Db 841 TGGCGCCATTACCGAGTCCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 900
QY 901 ACATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACATCAAAACAGGATTTTC 960
Db 901 ACATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACATCAAAACAGGATTTTC 960
QY 961 GCTGCTGGGGCAAAACAGCGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1020
Db 961 GCTGCTGGGGCAAAACAGCGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1020
QY 1021 AGGGCAATCAGTGTGTCCTGCTCACTGGTGAAGAAAGAAACCAACCTTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGTGTGTCCTGCTCACTGGTGAAGAAAGAAACCAACCTTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGGTGGCGGATTCATTAATGACGCTGGCACACAGGTTT 1140
```

Db 1081 |||||CGAAACCGCCTCTCCCCGCGGTTTGGCCGATTCATTAAATGAGCTGGCAGCAGGTTT1140
Qy 1141 CCCGACTCGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG1200
Db 1141 CCCGACTCGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG1200
Qy 1201 GCACAATTTCTGATGTTTGGAGCTTATCATGACGTGCGACGGTGCAACCAATGCTTCTGGCG1260
Db 1201 GCACAATTTCTGATGTTTGGAGCTTATCATGACGTGCGACGGTGCAACCAATGCTTCTGGCG1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTGATTAATCACTGCATTAATTCG1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTGATTAATCACTGCATTAATTCG1320
Qy 1321 TGTGCTCTCAAGCGCACTCCCGTTCCTGGATAATGTTTTTGGCGCGACATCATACGGTT1380
Db 1321 TGTGCTCTCAAGCGCACTCCCGTTCCTGGATAATGTTTTTGGCGCGACATCATACGGTT1380
Qy 1381 CTGGCAAAATTTCTGAAATGAGCTTTCACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTTTCACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA1440
Qy 1441 ATTTGAGCGGATACAAATTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA1500
Db 1441 ATTTGAGCGGATACAAATTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA1500
Qy 1501 GCACCTTTCACCAACAGGACCATAGATTAATGAAACTGGAAGAGTAAATCGTAAATCTGG1560
Db 1501 GCACCTTTCACCAACAGGACCATAGATTAATGAAACTGGAAGAGTAAATCGTAAATCTGG1560
Qy 1561 ATTTAAGCGGATAAAGGCTATAACGGTCTGCGTGAAGTCGGTAAGAAATTCGAGAAAGAT1620
Db 1561 ATTTAAGCGGATAAAGGCTATAACGGTCTGCGTGAAGTCGGTAAGAAATTCGAGAAAGAT1620
Qy 1621 ACCGGAATTAAGTCACCGTTGAGCATCCGGATAACTGGAGAGAAATTTCCACAGGTT1680
Db 1621 ACCGGAATTAAGTCACCGTTGAGCATCCGGATAACTGGAGAGAAATTTCCACAGGTT1680
Qy 1681 GCGGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC1740
Db 1681 GCGGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC1740
Qy 1741 GCTCAATCTGGCCTGTGGTGAATTCACCCGGACAAAGCGTTCCAGGACAAGCTGTAT1800
Db 1741 GCTCAATCTGGCCTGTGGTGAATTCACCCGGACAAAGCGTTCCAGGACAAGCTGTAT1800
Qy 1801 CCGTTTACCTGGGATGCGGTAGCTTACACGCAAGCTGATGCTTACCGATCGCTGTT1860
Db 1801 CCGTTTACCTGGGATGCGGTAGCTTACACGCAAGCTGATGCTTACCGATCGCTGTT1860
Qy 1861 GAAAGGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAACCTGGGAA1920
Db 1861 GAAAGGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAACCTGGGAA1920
Qy 1921 GAGATCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC1980
Db 1921 GAGATCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC1980
Qy 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG2040
Db 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG2040
Qy 2041 TATGAAACCGCAAGTACGACATTAAGACGTGGCGGTGGATAACGCTGGCGGAAAGCG2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGTGGCGGTGGATAACGCTGGCGGAAAGCG2100
Qy 2101 GGTCTGACCTTCTGTTGACCTGATTAATAAACAACATGAATGCGACACCGATTTAC2160
Db 2101 GGTCTGACCTTCTGTTGACCTGATTAATAAACAACATGAATGCGACACCGATTTAC2160
Qy 2161 TCCATTCGAGAGCTGCTTAAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG2220

Db 2161 TCCATTCGAGAGCTGCTTAAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG2220
Qy 2221 GCATGGTCCAAACATCGACACACGCAAAAGTGAATTAATGTGTAACGGTACTGCCGACCTTC2280
Db 2221 GCATGGTCCAAACATCGACACACGCAAAAGTGAATTAATGTGTAACGGTACTGCCGACCTTC2280
Qy 2281 AAGGTTCAACCATCCAAACCGTTTGGTGGCGTCTGAGCGCAGGTATTAACGCCGCCAGT2340
Db 2281 AAGGTTCAACCATCCAAACCGTTTGGTGGCGTCTGAGCGCAGGTATTAACGCCGCCAGT2340
Qy 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAACATATCTGCTGACTGATGAAGTCTG2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAACATATCTGCTGACTGATGAAGTCTG2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAAGTCTTTACGAGGAAG2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAAGTCTTTACGAGGAAG2460
Qy 2461 TTGGGAAAAGATTCACAGTATTTGGCGCCACATGGAACCGCCAGAAAGGTGAATCATG2520
Db 2461 TTGGGAAAAGATTCACAGTATTTGGCGCCACTATGGAACCGCCAGAAAGGTGAATCATG2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTATGCGTGCCTGCTGATCAACGCC2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTATGCGTGCCTGCTGATCAACGCC2580
Qy 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTTGAGCTCG2640
Db 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTTGAGCTCG2640
Qy 2641 AACAAACAACAATAACAATAACAACAACTCGGGATCGAGGAGGATTTGAGAAATTC2700
Db 2641 AACAAACAACAATAACAATAACAACAACTCGGGATCGAGGAGGATTTGAGAAATTC2700

RESULT 15
US-10-263-153-61
; Sequence 61, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 7370
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3447)
; OTHER INFORMATION: pMBP-c2X-Toxop30MIX1
US-10-263-153-61

Query Match 81.5%; Score 2690.4; DB 18; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGCGGTATGCGCATGATAGCGCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGCGGTATGCGCATGATAGCGCCGGAAGAGA 60
Qy 61 GTCAATTACGGGTGTTGAATGTAACCAACAGTAACTTATACATGTCGACAGTATGCGG 120
Db 61 GTCAATTACGGGTGTTGAATGTAACCAACAGTAACTTATACATGTCGACAGTATGCGG 120

QY	121	GTGTCCTCTTATCAGACCGTTTCCCGCTGGTGAACACGAGCGACGACGTTTCTGCGGAAA	180
Db	121	GTGTCCTCTTATCAGACCGTTTCCCGCTGGTGAACACGAGCGACGTTTCTGCGGAAA	180
QY	181	CGCGGAAAAGTGGAGCGCGGATGGCGAGCTGAATTAATTCACAAACCGGTGGCAC	240
Db	181	CGCGGAAAAGTGGAGCGCGGATGGCGAGCTGAATTAATTCACAAACCGGTGGCAC	240
QY	241	AACAACCTGGGCGGCAACAGCTGCTGATTCGCGCTGGCCACTCCAGTCTGGCCCTGC	300
Db	241	AACAACCTGGGCGGCAACAGCTGCTGATTCGCGCTGGCCACTCCAGTCTGGCCCTGC	300
QY	301	ACGGCGCGTGC3AAAATTGTCGGCGGATTAATACTTCGCGCGGATCAACTGGGTGCCAGCG	360
Db	301	ACGGCGCGTGC3AAAATTGTCGGCGGATTAATACTTCGCGCGGATCAACTGGGTGCCAGCG	360
QY	361	TGCTGTGTCGATGCTGATAGAACGAGCGGTGCAAGCCTGTAAGCGGCGGTGCAAAATC	420
Db	361	TGCTGTGTCGATGCTGATAGAACGAGCGGTGCAAGCCTGTAAGCGGCGGTGCAAAATC	420
QY	421	TTCTCGGCGAAACCGCTCAGTGGCTGATCAITTAATCTATCCGCTGGATGACACGAGGATGCCA	480
Db	421	TTCTCGGCGAAACCGCTCAGTGGCTGATCAITTAATCTATCCGCTGGATGACACGAGGATGCCA	480
QY	481	TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTGACCCAGA	540
Db	481	TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTGACCCAGA	540
QY	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGCTACGGCTGATGACACGAGGATGCCA	600
Db	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGCTACGGCTGATGACACGAGGATGCCA	600
QY	601	TGCTGCATTCGGTCAACAGCAAAATCGGCTGTTAGCGGCGCCATTAAGTTCGTCTCGG	660
Db	601	TGCTGCATTCGGTCAACAGCAAAATCGGCTGTTAGCGGCGCCATTAAGTTCGTCTCGG	660
QY	661	CGCGTCTGGCTCTGGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCGGATAG	720
Db	661	CGCGTCTGGCTCTGGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCGGATAG	720
QY	721	CGGAACGGGAAGGCGAGCTGAGTGCCATGTCCGCTTTTCAACAAACCATGCAAAATGCTGA	780
Db	721	CGGAACGGGAAGGCGAGCTGAGTGCCATGTCCGCTTTTCAACAAACCATGCAAAATGCTGA	780
QY	781	ATGAGGCGCATGTTCCCACTCGATGCTGGTGGCAACGATCAGATGGCGCTGGCGCAA	840
Db	781	ATGAGGCGCATGTTCCCACTCGATGCTGGTGGCAACGATCAGATGGCGCTGGCGCAA	840
QY	841	TGCGCGCATTAACGAGTCCGGCTGGCGTGGTGGGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGCGCATTAACGAGTCCGGCTGGCGTGGTGGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAAGACGCTCATGTTATATCCGCGGTTAAACCAATCAAAACAGGATTTTC	960
Db	901	ACGATACCGAAGACGCTCATGTTATATCCGCGGTTAAACCAATCAAAACAGGATTTTC	960
QY	961	GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020
Db	961	GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020
QY	1021	AGGGCAATCAGCTGTTGCCGTCTCACTGGTGAAGAAAACCAACCTCGGCGCCCAATA	1080
Db	1021	AGGGCAATCAGCTGTTGCCGTCTCACTGGTGAAGAAAACCAACCTCGGCGCCCAATA	1080
QY	1081	CGAAACCGGCTCTCCCGCGGTTGGCCGATTCATTATGAGCTGGACGACGAGTTT	1140
Db	1081	CGAAACCGGCTCTCCCGCGGTTGGCCGATTCATTATGAGCTGGACGACGAGTTT	1140
QY	1141	CCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTTAGTTAGTCTACTCATTTAG	1200
Db	1141	CCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTTAGTTAGTCTACTCATTTAG	1200

QY	1201	GCACAAATTTCTCATGTTTGACAGCTTATCATCGACTGCAACGCTGCACCAATCTCTTGCGG	1260
Db	1201	GCACAAATTTCTCATGTTTGACAGCTTATCATCGACTGCAACGCTGCACCAATCTCTTGCGG	1260
QY	1261	TCAGGACGCCATCCGAACTGCTGTGCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTG	1320
Db	1261	TCAGGACGCCATCCGAACTGCTGTGCTGAGCTGCTGAGCTGCTGAGCTGCTGAGCTGCTG	1320
QY	1321	TGTCGCTCAAGCGCGCACTCCCGTTCTGATAAATCTGATAAATCTGATAAATCTGATAAAT	1380
Db	1321	TGTCGCTCAAGCGCGCACTCCCGTTCTGATAAATCTGATAAATCTGATAAATCTGATAAAT	1380
QY	1381	CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA	1440
Db	1381	CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA	1440
QY	1441	ATTGTGAGCGGATAAACAAATTTTACACAGGAAACAGCCAGTCCGTTAGGTGTTTTCACGA	1500
Db	1441	ATTGTGAGCGGATAAACAAATTTTACACAGGAAACAGCCAGTCCGTTAGGTGTTTTCACGA	1500
QY	1501	GCATTCACCAACAAAGACCATAGATTTATGAAATCTGAAGAGTAAATCTGGTAAATCTGG	1560
Db	1501	GCATTCACCAACAAAGACCATAGATTTATGAAATCTGAAGAGTAAATCTGGTAAATCTGG	1560
QY	1561	ATTAAACGCGATAAAGCTATTAACGCTCTGCTGAACTGAACTGAACTGAACTGAACTGAA	1620
Db	1561	ATTAAACGCGATAAAGCTATTAACGCTCTGCTGAACTGAACTGAACTGAACTGAACTGAA	1620
QY	1621	ACCGAATTAAGCTACCGTTGAGCATCCGATTAACCTGGAAGAGTAAATCTGGAGTAAAT	1680
Db	1621	ACCGAATTAAGCTACCGTTGAGCATCCGATTAACCTGGAAGAGTAAATCTGGAGTAAAT	1680
QY	1681	GCAGCACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACAGACCGCTTGGTGGCTAC	1740
Db	1681	GCAGCACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACAGACCGCTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCTGTTGGCTGAAATCACCCGGAACAAAGCGTTCCAGGACAACTGTAAT	1800
Db	1741	GCTCAATCTGGCTGTTGGCTGAAATCACCCGGAACAAAGCGTTCCAGGACAACTGTAAT	1800
QY	1801	CGTTCCTGCGATGCGCTGATTAACAGCGCAAGCTGATTCCTGCGATCGCTGTT	1860
Db	1801	CGTTCCTGCGATGCGCTGATTAACAGCGCAAGCTGATTCCTGCGATCGCTGTT	1860
QY	1861	GAAAGCTTATCGCTGATTTATTAACAAAGCTGCTGCGGAAACCCGCAAAACCTGGGAA	1920
Db	1861	GAAAGCTTATCGCTGATTTATTAACAAAGCTGCTGCGGAAACCCGCAAAACCTGGGAA	1920
QY	1921	GAGATCCCGCGCTGGATAAGAACTGAAAGCGAAAGGTAAAGAGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATAAGAACTGAAAGCGAAAGGTAAAGAGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACGTAATTCACCTGGCGCTGATTCGCTGACCGGGGTTATGCGTTCAAG	2040
Db	1981	CTGCAAGAACGTAATTCACCTGGCGCTGATTCGCTGACCGGGGTTATGCGTTCAAG	2040
QY	2041	TATGAAAACGGCAAGTACGACATTAAGAGCTGGCGTGGATTAACGCTGGCGCGGAAAGCG	2100
Db	2041	TATGAAAACGGCAAGTACGACATTAAGAGCTGGCGTGGATTAACGCTGGCGCGGAAAGCG	2100
QY	2101	GCTCTGACCTTCTGGTTGACCTGATTAATAAAGGCGAAACAGGATGACCATCAACGCGCTTC	2160
Db	2101	GCTCTGACCTTCTGGTTGACCTGATTAATAAAGGCGAAACAGGATGACCATCAACGCGCTTC	2160
QY	2161	TCCATCGCAGAGCTGCTTTTAAAGGCGAAACAGGATGACCATCAACGCGCTTC	2220
Db	2161	TCCATCGCAGAGCTGCTTTTAAAGGCGAAACAGGATGACCATCAACGCGCTTC	2220
QY	2221	GCATGGTCCAACTGACACAGCAAAAGTGAATTAATGCTGTAACGCTGCTGCGGCTTC	2280
Db	2221	GCATGGTCCAACTGACACAGCAAAAGTGAATTAATGCTGTAACGCTGCTGCGGCTTC	2280
QY	2281	AAGGCTCAACCATCCAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAAAGCGCGCAGT	2340

D _b	2281	AAGGGTCAACCATCCAAACCGTTGCGTGCTGAGCGCAGGTATTAACGCCGCAGT	2340
Q _y	2341	CCGAACAAAGACTGGCAAAGAGTCCTCGAACAATTATCTGCTGACTGATGAAGGTCG	2400
D _b	2341	CCGAACAAAGACTGGCAAAGAGTCCTCGAACAATTATCTGCTGACTGATGAAGGTCG	2400
Q _y	2401	GAA CGGGTTAATAAGAACAACCCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAAG	2460
D _b	2401	GAA CGGGTTAATAAGAACAACCCTGGGTGCCGTAGCGCTGAAGTCTTACGAGGAAAG	2460
Q _y	2461	TTCGCGAAAGATCCACGTATTGCGCGCACCATGGAAAACGCCCAAGAAGGTGAATCATG	2520
D _b	2461	TTCGCGAAAGATCCACGTATTGCGCGCACCATGGAAAACGCCCAAGAAGGTGAATCATG	2520
Q _y	2521	CCGAACATCCGCGAGATGTCGCGTTTTCTGGTATCGCTGCGTACTCGGTTGATCAACGCC	2580
D _b	2521	CCGAACATCCGCGAGATGTCGCGTTTTCTGGTATCGCTGCGTACTCGGTTGATCAACGCC	2580
Q _y	2581	GCCAGCGGTGCTCAGACTGTGCGATGAAGCCCTGAAAGACGCCGCAGACTAATTCGAGTCTG	2640
D _b	2581	GCCAGCGGTGCTCAGACTGTGCGATGAAGCCCTGAAAGACGCCGCAGACTAATTCGAGTCTG	2640
Q _y	2641	AACAACAACAATAACAATAACAACCTCCGGATCGAGGGAAGGATTTGAAATTC	2700
D _b	2641	AACAACAACAATAACAATAACAACCTCCGGATCGAGGGAAGGATTTGAAATTC	2700

Search completed: August 3, 2005, 04:21:52
Job time : 1319.2 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 352.4 Seconds
(without alignments)
15322.682 Million cell updates/sec

Title: US-09-765-555B-15

Perfect score: 3300

Sequence: 1 ccgacacacgaatgtgc.....acgacgtccggactacgct 3300

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 81813859 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description	
1	2700	81.8	7475	2	US-08-971-036-1	
2	2700	81.8	7475	3	US-09-096-570-1	
3	2700	81.8	7475	3	US-09-265-617B-1	
C	4	1417.6	43.0	5926	3	US-09-027-169-3
	5	1362	41.3	5201	4	US-09-640-882-2
	6	1362	41.3	5201	4	US-09-640-882-3
	7	1204	36.5	4557	4	US-08-778-717-5
8	1200.8	36.4	3832	1	US-08-148-675A-2	
C	9	1199.2	36.3	5248	3	US-08-487-283A-18
	10	1199.2	36.3	5248	5	PCT-US96-05611A-21
C	11	1199.2	36.3	5312	4	US-10-263-103-35
C	12	1199.2	36.3	5443	2	US-08-929-967-1
C	13	1199.2	36.3	5502	4	US-09-702-705-785
C	14	1199.2	36.3	5502	4	US-09-736-457-785
C	15	1199.2	36.3	5502	4	US-09-614-124B-785
C	16	1199.2	36.3	5502	4	US-09-671-325-785
C	17	1199.2	36.3	5502	4	US-09-589-184-785
C	18	1199.2	36.3	5502	4	US-09-658-824-785
C	19	1199.2	36.3	5616	2	US-08-929-967-3
C	20	1199.2	36.3	5873	4	US-09-695-437A-62
C	21	1199.2	36.3	6353	4	US-09-702-705-784
C	22	1199.2	36.3	6353	4	US-09-736-457-784
C	23	1199.2	36.3	6353	4	US-09-614-124B-784
C	24	1199.2	36.3	6353	4	US-09-671-325-784
C	25	1199.2	36.3	6353	4	US-09-589-184-784
C	26	1199.2	36.3	6353	4	US-09-658-824-784
	27	1199.2	36.3	6361	3	US-08-646-538-7
					Sequence 1, Appli	
					Sequence 1, Appli	
					Sequence 1, Appli	
					Sequence 2, Appli	
					Sequence 3, Appli	
					Sequence 5, Appli	
					Sequence 2, Appli	
					Sequence 18, Appli	
					Sequence 21, Appli	
					Sequence 35, Appli	
					Sequence 1, Appli	
					Sequence 785, App	
					Sequence 785, App	
					Sequence 785, App	
					Sequence 785, App	
					Sequence 785, App	
					Sequence 62, Appli	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	
					Sequence 784, App	

C	28	1199.2	36.3	6361	3	US-09-503-222-7
C	29	1199.2	36.3	6727	2	US-08-125-462-2
C	30	1199.2	36.3	6727	2	US-08-891-848-2
C	31	1199.2	36.3	6799	2	US-08-125-462-5
C	32	1199.2	36.3	6799	2	US-08-891-848-5
C	33	1199.2	36.3	7659	3	US-09-128-314-4
C	34	1199.2	36.3	7659	3	US-09-056-556-213
C	35	1199.2	36.3	7676	3	US-03-072-596-208
C	36	1199.2	36.3	7676	4	US-09-072-967-213
C	37	1199.2	36.3	7676	4	US-09-287-849-9
C	38	1199.2	36.3	8031	3	US-09-643-597-254
C	39	1199.2	36.3	8031	4	US-09-480-884A-254
C	40	1199.2	36.3	8031	4	US-09-542-615A-254
C	41	1199.2	36.3	8031	4	US-03-606-421B-254
C	42	1199.2	36.3	8031	4	US-09-476-496A-254
C	43	1199.2	36.3	8031	4	US-09-630-940B-254
C	44	1199.2	36.3	8157	3	US-09-128-314-3
C	45	1199.2	36.3	8501	3	US-08-793-900-1

ALIGNMENTS

RESULT 1
US-08-971-036-1
; Sequence 1, Application US/08971036
; Patent No. 5866684
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R
; APPLICANT: Hurst, David N
; APPLICANT: Jones, Philip S
; APPLICANT: Kay, Paul B
; APPLICANT: Raynham, Tony M
; APPLICANT: Wilson, Francis X
; TITLE OF INVENTION: Amino Acid Derivatives
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: N.J.
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/971.036
; FILING DATE: 14-NOV-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: GB 9623908.2
; FILING DATE: 18-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kreisler, Lewis J
; REGISTRATION NUMBER: 38522
; REFERENCE/DOCKET NUMBER: RAN 4430/073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (973) 235-4387
; TELEFAX: (973) 235-2163
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7475 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
US-08-971-036-1

Query Match 81.8%; Score 2700; DB 2; Length 7475;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	CGGACACCATGAAATGGTGCAGAAACCTTTCCGGGTATGCGCATGATAGCGCCGGAGAGA	60
Db	1	CGGACACCATGAAATGGTGCAGAAACCTTTCCGGGTATGCGCATGATAGCGCCGGAGAGA	60
QY	61	GTCAATTGAGGGTGGTGAATGTGAACACAGTAAACGTTATACGATCTCCAGAGTATCGG	120
Db	61	GTCAATTGAGGGTGGTGAATGTGAACACAGTAAACGTTATACGATCTCCAGAGTATCGG	120
QY	121	GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACACGCGCAGCCAGCTTTCTCGAAAA	180
Db	121	GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACACGCGCAGCCAGCTTTCTCGAAAA	180
QY	181	CGCGGAAAAAGTGAAGCGGATGCGGAGCTGAATATATCCCAACCGCTGGCAC	240
Db	181	CGCGGAAAAAGTGAAGCGGATGCGGAGCTGAATATATCCCAACCGCTGGCAC	240
QY	241	AACAACTGGCGGCAACAGTCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
Db	241	AACAACTGGCGGCAACAGTCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
QY	301	ACGGCGCTCCCAATTTGTCGGGGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360
Db	301	ACGGCGCTCCCAATTTGTCGGGGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360
QY	361	TGTTGGTGTCCGATGATAGAACGAGCGGCTGGAAGCCTGTAAGCGCGGTGCACAATC	420
Db	361	TGTTGGTGTCCGATGATAGAACGAGCGGCTGGAAGCCTGTAAGCGCGGTGCACAATC	420
QY	421	TTCTCGCGCAACGGCTCAGTGGGTGATCAATTAATCTCCGCTGGATGACCAAGATGCCA	480
Db	421	TTCTCGCGCAACGGCTCAGTGGGTGATCAATTAATCTCCGCTGGATGACCAAGATGCCA	480
QY	481	TTGCTGTGGAAGCTGCGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACC	540
Db	481	TTGCTGTGGAAGCTGCGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACC	540
QY	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGGCTGGGGTGGAGATC	600
Db	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGGCTGGGGTGGAGATC	600
QY	601	TGTTGCGATTTGGTTCACAGCAATTCGCGTGTAGCGGGCCATTAAGTTCTGTTCTCGG	660
Db	601	TGTTGCGATTTGGTTCACAGCAATTCGCGTGTAGCGGGCCATTAAGTTCTGTTCTCGG	660
QY	661	CGGCTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAATTCAGCCGATAG	720
Db	661	CGGCTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAATTCAGCCGATAG	720
QY	721	CGGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTCACAAACCATGCAATGCTGA	780
Db	721	CGGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTTCACAAACCATGCAATGCTGA	780
QY	781	ATGAGGCGATGTTCCCACTGCGATGCTGTTGCGCAACGATCAGATGGCGCTGGCGCAA	840
Db	781	ATGAGGCGATGTTCCCACTGCGATGCTGTTGCGCAACGATCAGATGGCGCTGGCGCAA	840
QY	841	TGCGCGCATTTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGCGCATTTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAAGACAGTCTATGTTATATCCCGGTTTAAACCATCAACAGGATTTTC	960
Db	901	ACGATACCGAAGACAGTCTATGTTATATCCCGGTTTAAACCATCAACAGGATTTTC	960
QY	961	GCCTGCTGGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCAGCGGTGA	1020
Db	961	GCCTGCTGGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCAGCGGTGA	1020
QY	1021	AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAAACACCTGGCCCAATA	1080
Db	1021	AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAAACACCTGGCCCAATA	1080

Db	1021	AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAAACACCTGGCCCAATA	1080
QY	1081	CGAAAAACCGCTCTCCCGCGGCTTGGCCGATTCATTAATGACAGTGGCACAGAGGTTT	1140
Db	1081	CGAAAAACCGCTCTCCCGCGGCTTGGCCGATTCATTAATGACAGTGGCACAGAGGTTT	1140
QY	1141	CCGACCTGGAAGAGGGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCATTATAG	1200
Db	1141	CCGACCTGGAAGAGGGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCATTATAG	1200
QY	1201	GCACAAATCTCATGTTTGCACAGCTTATCATCGACTGACAGCTGACCAATGCTTCTCGCG	1260
Db	1201	GCACAAATCTCATGTTTGCACAGCTTATCATCGACTGACAGCTGACCAATGCTTCTCGCG	1260
QY	1261	TCAGGACGCCATCGAAAGCTGTGGTATGCTGTGACAGGTCTGTAATCACTGTCATATTCG	1320
Db	1261	TCAGGACGCCATCGAAAGCTGTGGTATGCTGTGACAGGTCTGTAATCACTGTCATATTCG	1320
QY	1321	TGTCGCTCAAGGCGCACTCCCGTCTCGGTAATGTTTTTGGCCGCGACATCAACCGTT	1380
Db	1321	TGTCGCTCAAGGCGCACTCCCGTCTCGGTAATGTTTTTGGCCGCGACATCAACCGTT	1380
QY	1381	CTGCAAAATTTCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTAAATGTTGCGCG	1440
Db	1381	CTGCAAAATTTCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTAAATGTTGCGCG	1440
QY	1441	ATTGTGACCGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCA	1500
Db	1441	ATTGTGACCGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCA	1500
QY	1501	GCATTTCAACCAACAGGACCATAGATTATGAAATCTGAAAGGTAACTGTTAACTCTGG	1560
Db	1501	GCATTTCAACCAACAGGACCATAGATTATGAAATCTGAAAGGTAACTGTTAACTCTGG	1560
QY	1561	ATTAAACCGGATTAAGGCTTAAACCGTCTCGTGAAGTGGTAAAGAAATTCGAGAAAG	1620
Db	1561	ATTAAACCGGATTAAGGCTTAAACCGTCTCGTGAAGTGGTAAAGAAATTCGAGAAAG	1620
QY	1621	ACCGGAATTAAGCTACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCGCCACAG	1680
Db	1621	ACCGGAATTAAGCTACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCGCCACAG	1680
QY	1681	CGCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCAACACCGCTTTGGTGGCTAC	1740
Db	1681	CGCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCAACACCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGAT	1800
Db	1741	GCTCAATCTGGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGAT	1800
QY	1801	CCGTTTACCTGGGATCGCGTACGTTACACGGCAAGCTGATTCGTTACCGGATCGCTGTT	1860
Db	1801	CCGTTTACCTGGGATCGCGTACGTTACACGGCAAGCTGATTCGTTACCGGATCGCTGTT	1860
QY	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTCCGAAACCCGCAAAACCTCGGAA	1920
Db	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTCCGAAACCCGCAAAACCTCGGAA	1920
QY	1921	GAGATCCCGCGCTGATTAAGAACTGAAAGGTAAAGAGCGCGCTGATGTTTCAAC	1980
Db	1921	GAGATCCCGCGCTGATTAAGAACTGAAAGGTAAAGAGCGCGCTGATGTTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGGGTTATGGTTCAG	2040
Db	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGGGTTATGGTTCAG	2040
QY	2041	TATGAAAAACCGCAAGTACGCAATTAAGACGTGGCGTGGATTAACGCTGGCGGAAAGCG	2100
Db	2041	TATGAAAAACCGCAAGTACGCAATTAAGACGTGGCGTGGATTAACGCTGGCGGAAAGCG	2100
QY	2101	GGTCTGACCTTCTGTTGACCTGATTAAGAAACCAACACATGACACACCGATTAC	2160
Db	2101	GGTCTGACCTTCTGTTGACCTGATTAAGAAACCAACATGAATGACACACCGATTAC	2160

QY 841 TGCGGCCATTACGAGTCCGGCTCGCGTTGGTGGGATATCTCGTAGTGGATACG 900
DB 841 TGCGGCCATTACGAGTCCGGCTCGCGTTGGTGGGATATCTCGTAGTGGATACG 900
QY 901 ACGATACCGAAGACGAGCTCATGTTATATCCCGCGTAAACACAGGATTTTC 960
DB 901 ACGATACCGAAGACGAGCTCATGTTATATCCCGCGTAAACACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACGAGCTGAGCCGTTGCTGCACTCTCTCAGGCCAGCGGTGA 1020
DB 961 GCCTGCTGGGCAACACGAGCTGAGCCGTTGCTGCACTCTCTCAGGCCAGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTGGCCGCTCTCACTGTGAAAGAAACCAACCCCTGGCGCCCAATA 1080
DB 1021 AGGCAATCAGCTGTGGCCGCTCTCACTGTGAAAGAAACCAACCCCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCAATTAATGCAGCTGGCAGCAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCAATTAATGCAGCTGGCAGCAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGTAGTTAGCTCAGTCATTAG 1200
DB 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGTAGTTAGCTCAGTCATTAG 1200
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260
DB 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGAGCCATCGAAGCTGTGATGCTGTGAGTATGCTTTCGCGCGCAGCATCAACGGTT 1320
DB 1261 TCAGGAGCCATCGAAGCTGTGATGCTGTGAGTATGCTTTCGCGCGCAGCATCAACGGTT 1320
QY 1321 TGTGCTCAAGCGCAGCTCCCGTTCTGGATTAATGTTTTCGCGCGCAGCATCAACGGTT 1380
DB 1321 TGTGCTCAAGCGCAGCTCCCGTTCTGGATTAATGTTTTCGCGCGCAGCATCAACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
DB 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
QY 1441 ATTGTAGCGGATTAACAAATTTTCAACAGGAAACAGCAGTCCGTTTATAGTTTTCACGA 1500
DB 1441 ATTGTAGCGGATTAACAAATTTTCAACAGGAAACAGCAGTCCGTTTATAGTTTTCACGA 1500
QY 1501 GCACCTTCAACAAAGGACCATAGATTAATGAAATCTGAAGTGAAGTAACTGCTAACTCG 1560
DB 1501 GCACCTTCAACAAAGGACCATAGATTAATGAAATCTGAAGTGAAGTAACTGCTAACTCG 1560
QY 1561 ATTAACCGCGATTAAGGCTATAACGGTCTCGCTGAAGTTCGTAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTAACCGCGATTAAGGCTATAACGGTCTCGCTGAAGTTCGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTACCGTTGAGCATCCGGATTAATCTGGAAGAGAAATTCGCCACAGTT 1680
DB 1621 ACCGGAATTAAGTACCGTTGAGCATCCGGATTAATCTGGAAGAGAAATTCGCCACAGTT 1680
QY 1681 GCGCAACTGCGATGCGCTGACATTAATCTTCTGGGCACACGACCGCTTGGTGGCTAC 1740
DB 1681 GCGCAACTGCGATGCGCTGACATTAATCTTCTGGGCACACGACCGCTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
DB 1741 GCTCAATCTGCGCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGGAAGCTGATGCTTACCGGATCGCTGTT 1860
DB 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGGAAGCTGATGCTTACCGGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCGCCCAAAACCTCGGAA 1920
DB 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCGCCCAAAACCTCGGAA 1920

QY 1921 GAGATCCGCGCTGGATATAAGAACTCAAGCGAAAGGTAAAGCGCGCTGATCTTCAAC 1980
DB 1921 GAGATCCGCGCTGGATATAAGAACTCAAGCGAAAGGTAAAGCGCGCTGATCTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTTGCTGCTGACGGGGTATGCGCTTCAAG 2040
DB 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTTGCTGCTGACGGGGTATGCGCTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTAAGACGTTGGCGCTGGATTAACGCTGGCGCAAGCG 2100
DB 2041 TATGAAACCGCAAGTACGACATTAAGACGTTGGCGCTGGATTAACGCTGGCGCAAGCG 2100
QY 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAAACCAACACATGAATGACAGACACCGATTAC 2160
DB 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAAACCAACACATGAATGACAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAAGCTGCTTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
DB 2161 TCCATCGCAGAAGCTGCTTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGTTCCAACTGACACACAGCAAGTGAATTAATGCTGTAACGGTACTGCGCACCTTC 2280
DB 2221 GCATGTTCCAACTGACACACAGCAAGTGAATTAATGCTGTAACGGTACTGCGCACCTTC 2280
QY 2281 AAGGTTCAACATCTCAACCGCTTCTGTTGGCTGCTGAGCGCAGGTATTAACGGCCCGTGG 2340
DB 2281 AAGGTTCAACATCTCAACCGCTTCTGTTGGCTGCTGAGCGCAGGTATTAACGGCCCGTGG 2340
QY 2341 CCAGAACAAAGAGCTGGCAAGAGTCTCTCGAAAACTATCTGCTGACTGATGAAGGCTCTG 2400
DB 2341 CCAGAACAAAGAGCTGGCAAGAGTCTCTCGAAAACTATCTGCTGACTGATGAAGGCTCTG 2400
QY 2401 GAAGCGTTTAATAAGCAAAACCGCTGGTGGCTGAGCGTGAAGTCTTACGAGGAAGAG 2460
DB 2401 GAAGCGTTTAATAAGCAAAACCGCTGGTGGCTGAGCGTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCACGTATTGCGCCACCATCGAAACCGCCAGAAAGGTGAAATCATG 2520
DB 2461 TTGGCGAAAGATCCACGTATTGCGCCACCATCGAAACCGCCAGAAAGGTGAAATCATG 2520
QY 2521 CCAGAACATCCGCGAGATGCTCCGCTTTCTGCTGATGCGGTGCTGCTGATGATCAACGCC 2580
DB 2521 CCAGAACATCCGCGAGATGCTCCGCTTTCTGCTGATGCGGTGCTGCTGATGATCAACGCC 2580
QY 2581 GCCAGCGCTGCTCAGACTGCTGATGAGCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
DB 2581 GCCAGCGCTGCTCAGACTGCTGATGAGCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAAATAACAAATAACAAATAACAAATAACAAATAACAAATAACAAATAAC 2700
DB 2641 AACAAACAAACAAATAACAAATAACAAATAACAAATAACAAATAACAAATAACAAATAAC 2700

RESULT 3

US-09-265-617B-1
; Sequence 1, Application US/09265617B
; Patent No. 637283
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R.
; APPLICANT: Hurst, David N.
; APPLICANT: Jones, Philip S.
; APPLICANT: Kay, Paul B.
; APPLICANT: Ravnham, Tony M.
; APPLICANT: Wilson, Francis X.
; TITLE OF INVENTION: Antiviral Medicaments
; FILE REFERENCE: 20052 antiviral medicaments
; CURRENT APPLICATION NUMBER: US/09/265,617B
; CURRENT FILING DATE: 1999-03-10
; PRIOR APPLICATION NUMBER: GB9806815.8
; PRIOR FILING DATE: 1998-03-30
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1

; LENGTH: 7475									
; TYPE: DNA									
; ORGANISM: Artificial Sequence									
; FEATURE:									
; OTHER INFORMATION: Description of Artificial Sequence: Vector and									
; OTHER INFORMATION: Gene Fragments									
US-09-265-617B-1									
Query Match 81.8%; Score 2700; DB 3; Length 7475;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	CCGACACCATCAATGGTGC	AAACCTTTCCGGGTATG	CGCATGATAGCCCGGAAGAGA	60				
Db	1	CCGACACCATCAATGGTGC	AAACCTTTCCGGGTATG	CGCATGATAGCCCGGAAGAGA	60				
Qy	61	GTCATTCAGGGTGGTGAAT	GTGAACACAGTAACGTT	ATACGATGTCGACAGTATGCCG	120				
Db	61	GTCATTCAGGGTGGTGAAT	GTGAACACAGTAAACGTT	ATACGATGTCGACAGTATGCCG	120				
Qy	121	GTGTCCTTATCAGACCGG	TTTCCCGGTGGTGAAC	CCAGGCCAGCCAGCGTTTCTGCG	180				
Db	121	GTGTCCTTATCAGACCGG	TTTCCCGGTGGTGAAC	CCAGGCCAGCCAGCGTTTCTGCG	180				
Qy	181	CGCGGAAATAGTGAACG	CGCGATGGCGGAGCTGA	ATTACATTCCTCCACCGCGTGGC	240				
Db	181	CGCGGAAATAGTGAACG	CGCGATGGCGGAGCTGA	ATTACATTCCTCCACCGCGTGGC	240				
Qy	241	AACAACTGGCGGCAAA	CAGTCGTTGCTGATGG	CGGTTGCCACCTCCAGTCTGGCCCTGC	300				
Db	241	AACAACTGGCGGCAAA	CAGTCGTTGCTGATGG	CGGTTGCCACCTCCAGTCTGGCCCTGC	300				
Qy	301	AGCGCGGTGCGCAAT	TGTCGGCGGATPAAAT	CTCGCGCGGATCAA	CTGGGTGGCCAGCG	360			
Db	301	AGCGCGGTGCGCAAT	TGTCGGCGGATPAAAT	CTCGCGCGGATCAA	CTGGGTGGCCAGCG	360			
Qy	361	TGGTGGTTCGATGTAG	AAACGAGCGGCTGGA	AGCCTGTAAAGCGCGGTGCA	420				
Db	361	TGGTGGTTCGATGTAG	AAACGAGCGGCTGGA	AGCCTGTAAAGCGCGGTGCA	420				
Qy	421	TTCTCGCGCAACGCT	CAGTGGGCTGATCAT	TAACTATCCGCTGGATGAC	CAGGATGCCA	480			
Db	421	TTCTCGCGCAACGCT	CAGTGGGCTGATCAT	TAACTATCCGCTGGATGAC	CAGGATGCCA	480			
Qy	481	TTGCTGTGGAAGCT	CGCTGCCTAAATGTT	CCGCGGTTATTTCT	TGATGTCCTGAC	540			
Db	481	TTGCTGTGGAAGCT	CGCTGCCTAAATGTT	CCGCGGTTATTTCT	TGATGTCCTGAC	540			
Qy	541	CACCCATCAACAGTA	TATTTTCTCCATGA	AGACGGTACGGACTGG	CGGTGGAGCATC	600			
Db	541	CACCCATCAACAGTA	TATTTTCTCCATGA	AGACGGTACGGACTGG	CGGTGGAGCATC	600			
Qy	601	TGGTCGCAATGGGT	CACAGCAAACTCG	CTGTAGCGGCCCATT	TAAGTCTGTCTCGG	660			
Db	601	TGGTCGCAATGGGT	CACAGCAAACTCG	CTGTAGCGGCCCATT	TAAGTCTGTCTCGG	660			
Qy	661	CGCGTCTGCGTCTG	GTGGCTGGCAATAAT	CTCACTCGCAATCAA	ATTACGCGGATAG	720			
Db	661	CGCGTCTGCGTCTG	GTGGCTGGCAATAAT	CTCACTCGCAATCAA	ATTACGCGGATAG	720			
Qy	721	CGGAACGGGAAGCG	AGCTGGAGTGCGAT	GTCCGGTTTTCA	AAACCATGCAAAATGCTGA	780			
Db	721	CGGAACGGGAAGCG	AGCTGGAGTGCGAT	GTCCGGTTTTCA	AAACCATGCAAAATGCTGA	780			
Qy	781	ATGAGGGCATCGT	CCCTCACTCGAGT	CGTGGTGGCCAA	CCGATCAGATGGCGCTGGCGCA	840			
Db	781	ATGAGGGCATCGT	CCCTCACTCGAGT	CGTGGTGGCCAA	CCGATCAGATGGCGCTGGCGCA	840			
Qy	841	TGCGGCCATTAC	CGAGTCCGGGCTG	CGGTTGGTGGGAT	ATCTCGGTAGTGGGATAG	900			
Db	841	TGCGGCCATTAC	CGAGTCCGGGCTG	CGGTTGGTGGGAT	ATCTCGGTAGTGGGATAG	900			
Qy	901	ACGATACCGAAG	ACAGCTCATGTTTA	TATCCCGCGGTTAA	CCACCATCAAAAGGATTTTC	960			

Db	901	ACGATACCGAAGACAGCT	CAATGTTATATCCCGCG	GTAAACCAACATCAAC	AGGATTTTC	960			
Qy	961	GCCTCTGGGGCAAA	CCAGCGTGGACCGCT	TTGCTGCAACTCTCT	CAGGCGCGGTGA	1020			
Db	961	GCCTCTGGGGCAAA	CCAGCGTGGACCGCT	TTGCTGCAACTCTCT	CAGGCGCGGTGA	1020			
Qy	1021	AGGGCAATCAGCT	GTGGCCGCTCTCACT	GTGGTGAAGAAACCA	CCCTGGCGCCCAATA	1080			
Db	1021	AGGGCAATCAGCT	GTGGCCGCTCTCACT	GTGGTGAAGAAACCA	CCCTGGCGCCCAATA	1080			
Qy	1081	CGAAAAACCGCT	CTCCCGCGCTTGGCG	GAATCAATTAATGC	AGTGGACGAGGTTT	1140			
Db	1081	CGAAAAACCGCT	CTCCCGCGCTTGGCG	GAATCAATTAATGC	AGTGGACGAGGTTT	1140			
Qy	1141	CCGCACTGGAAG	CGGGCAGTGAGCG	CAACCAATTAATGT	AGTGTAGCTCACTCAT	1200			
Db	1141	CCGCACTGGAAG	CGGGCAGTGAGCG	CAACCAATTAATGT	AGTGTAGCTCACTCAT	1200			
Qy	1201	GCACAAATTCAT	GTGTGACAGCTTAT	CATGCATGCGGTG	CAACCAATGCTTCTGGG	1260			
Db	1201	GCACAAATTCAT	GTGTGACAGCTTAT	CATGCATGCGGTG	CAACCAATGCTTCTGGG	1260			
Qy	1261	TCAGCAGCCAT	TCGGAAGCTGTG	GTATGCTGTGAG	GTGTAATCACTGCAT	1320			
Db	1261	TCAGCAGCCAT	TCGGAAGCTGTG	GTATGCTGTGAG	GTGTAATCACTGCAT	1320			
Qy	1321	TGTCGCTCAAG	CGCGCCTCCCGT	TCGTAATGTTTTT	TGCGCGGACATCA	1380			
Db	1321	TGTCGCTCAAG	CGCGCCTCCCGT	TCGTAATGTTTTT	TGCGCGGACATCA	1380			
Qy	1381	CTGGCAAAATAT	CTGAAATGAGCT	GTGCAATTAATCAT	CGGCTCGTAAATG	1440			
Db	1381	CTGGCAAAATAT	CTGAAATGAGCT	GTGCAATTAATCAT	CGGCTCGTAAATG	1440			
Qy	1441	ATTGTGACGGAT	AAACAAATTTCA	CACAGAAACGCCAG	TCCGTTTAGTGTTT	1500			
Db	1441	ATTGTGACGGAT	AAACAAATTTCA	CACAGAAACGCCAG	TCCGTTTAGTGTTT	1500			
Qy	1501	GCACCTTCAAC	CAAGGACCATAGA	TTATGAAAACTGA	AGAGGTAATCTGG	1560			
Db	1501	GCACCTTCAAC	CAAGGACCATAGA	TTATGAAAACTGA	AGAGGTAATCTGG	1560			
Qy	1561	ATTAACCGCGAT	AAAGGCTATAA	CGGTCTCGCTGA	AGTCGGTAAGAAAT	1620			
Db	1561	ATTAACCGCGAT	AAAGGCTATAA	CGGTCTCGCTGA	AGTCGGTAAGAAAT	1620			
Qy	1621	ACCGGAATTAAG	TACACCGTTCG	GAATAACTGG	AGAGAAATTC	1680			
Db	1621	ACCGGAATTAAG	TACACCGTTCG	GAATAACTGG	AGAGAAATTC	1680			
Qy	1681	CGCGCAACTG	CGATGGCCCTG	ACATTAATCTT	CGGCAACAGCGCT	1740			
Db	1681	CGCGCAACTG	CGATGGCCCTG	ACATTAATCTT	CGGCAACAGCGCT	1740			
Qy	1741	GCTCAATCTG	CGCTGTGGCTG	AAATCAACCCG	GAAGCGTTCAG	1800			
Db	1741	GCTCAATCTG	CGCTGTGGCTG	AAATCAACCCG	GAAGCGTTCAG	1800			
Qy	1801	CCGTTTACCT	GGGATGCGGTAG	CTTACAAACG	CAAGCTGATTC	1860			
Db	1801	CCGTTTACCT	GGGATGCGGTAG	CTTACAAACG	CAAGCTGATTC	1860			
Qy	1861	GAAGGTTTAT	CGCTGATTTTAA	CAAGATCTG	CTGCGCAACCC	1920			
Db	1861	GAAGGTTTAT	CGCTGATTTTAA	CAAGATCTG	CTGCGCAACCC	1920			
Qy	1921	GAGATCCCG	CGCTGGATAAAG	AACTGAAGCG	AAAGGTAAGAG	1980			
Db	1921	GAGATCCCG	CGCTGGATAAAG	AACTGAAGCG	AAAGGTAAGAG	1980			
Qy	1981	CTGCAAGAAC	CGTACTTCACT	CGCGCTGAT	TGCTGTGACG	2040			

```
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGACGGGGTATTCGCTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTAAGACCTGGCGTGGATACGCTGGCGCAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACCTGGCGTGGATACGCTGGCGCAAGCG 2100
Qy 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAATAAACAACAACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAATAAACAACAACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCCGAGAGCTGCTTAAATAGAGCGCAACAGCGATGACCATCAACGGCCGCTGG 2220
Db 2161 TCCATCCGAGAGCTGCTTAAATAGAGCGCAACAGCGATGACCATCAACGGCCGCTGG 2220
Qy 2221 GCATGGTCCAAATCGACACCAAGTGAATTAATGCTGTAACCGTACTCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACCAAGTGAATTAATGCTGTAACCGTACTCCGACCTTC 2280
Qy 2281 AAGGTCACCATCAACACCGTTCGTTGGCTGCTGAGCGCAGGTATTAACGCCCGCAGT 2340
Db 2281 AAGGTCACCATCAACACCGTTCGTTGGCTGCTGAGCGCAGGTATTAACGCCCGCAGT 2340
Qy 2341 CCGAACCAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACCAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGGTAAATAAGACCAACCGCTGGTGCCTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTAAATAAGACCAACCGCTGGTGCCTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGCGGAAAGATCCAGTATTTGCGCGCACCATGAAAGCCGCAAGAGGTGAATCATG 2520
Db 2461 TTGCGGAAAGATCCAGTATTTGCGCGCACCATGAAAGCCGCAAGAGGTGAATCATG 2520
Qy 2521 CCGAACATCCGCGAGATGCTCGCTTTCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCGCGAGATGCTCGCTTTCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Qy 2581 GCCAGCGTCTGACAGTCTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTCTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGATTTCAAGATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGATTTCAAGATTC 2700

RESULT 4
US-09-027-169-3/c
; Sequence 3, Application US/09027169
; Patent No. 6420524
; GENERAL INFORMATION:
; APPLICANT: CRAIG, NANCY L
; TITLE OF INVENTION: GAIN OF FUNCTION MUTATIONS IN
; TITLE OF INVENTION: ATP-DEPENDENT TRANSCRIPTION PROTEINS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESS: Anne Brown (Alston & Bird, LLP)
; STREET: 3605 Glenwood Ave.
; CITY: Raleigh
; STATE: NC
; COUNTRY: USA
; ZIP: 27608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/027,169
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Brown, Anne
; REGISTRATION NUMBER: 36,463
; REFERENCE/DOCKET NUMBER: 5789-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919 420 2205
; TELEFAX: 919 881 3175
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5926 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "pEM delta R.adj to 1"
; US-09-027-169-3

Query Match 43.0%; Score 1417.6; DB 3; Length 5926;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1455; Conservative 0; Mismatches 9; Indels 12; Gaps 2;

Qy 3 GACACCATCGAATGGTGCRAAAGCTTTTCGGGTATGCGCATGATAGCGCCGGAAGAGT 62
Db 1564 GACACCATCGAATGGTGCRAAAGCTTTTCGGGTATGCGCATGATAGCGCCGGAAGAGT 1505
Qy 63 CAATTTCAGGGTGGTGAATGTGAACACCAAGTAAACGTTATACGATGTCGAGAGTATGCCGGT 122
Db 1504 CAATTTCAGGGTGGTGAATGTGAACACCAAGTAAACGTTATACGATGTCGAGAGTATGCCGGT 1445
Qy 123 GTCTCTTATCAGACCGTTTCCCGGGTGTGAACAGGCCAGCCACGTTTCTGGAAAAACG 182
Db 1444 GTCTCTTATCAGACCGTTTCCCGGGTGTGAACAGGCCAGCCACGTTTCTGGAAAAACG 1385
Qy 183 CGGGAAGAGTGAAGCGGCGATGGCGGAGCTCAATTTACATTTCCACACCGGTGGCAAA 242
Db 1384 CGGGAAGAGTGAAGCGGCGATGGCGGAGCTCAATTTACATTTCCACACCGGTGGCAAA 1325
Qy 243 CAATCGCGCGCAACAGTCTGTTGCTGATTTGGGGTTGCCACCTCCAGTCTGGGCCCTGCAC 302
Db 1324 CAATCGCGCGCAACAGTCTGTTGCTGATTTGGGGTTGCCACCTCCAGTCTGGGCCCTGCAC 1265
Qy 303 GCSCGTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCGT 362
Db 1264 GCSCGTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCGT 1205
Qy 363 GTGCTGTCGATGTAGAAGCAAGCGCGTGAAGCGCTGTAAAGCGGCGGTGCAATCTT 422
Db 1204 GTGCTGTCGATGTAGAAGCAAGCGCGTGAAGCGCTGTAAAGCGGCGGTGCAATCTT 1145
Qy 423 CTCGCGCAACGCTCAGTGGGCTGATCATTAACATATCCGCTGGATGACCAAGGATGCCATT 482
Db 1144 CTCGCGCAACGCTCAGTGGGCTGATCATTAACATATCCGCTGGATGACCAAGGATGCCATT 1085
Qy 483 GCTGTGGAAGCTGCTGCATTAATGTTCCGGCGTTATTTCTTGATGTTCTTGACCAAGACA 542
Db 1084 GCTGTGGAAGCTGCTGCATTAATGTTCCGGCGTTATTTCTTGATGTTCTTGACCAAGACA 1025
Qy 543 CCCATCAACAGTATTTATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATCTG 602
Db 1024 CCCATCAACAGTATTTATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATCTG 965
Qy 603 GTCGCAATGGGTCAACAGCAAAATCGCGCTGTATAGCGGGCCCATTAAGTTCTGTCTGGCG 662
Db 964 GTCGCAATGGGTCAACAGCAAAATCGCGCTGTATAGCGGGCCCATTAAGTTCTGTCTGGCG 905
Qy 663 CGTCTGGTCTGGCTGGCGATTAATATCTCACTCGCAATCAAAATTCAGCGGATAGCG 722
Db 904 CGTCTGGTCTGGCTGGCGATTAATATCTCACTCGCAATCAAAATTCAGCGGATAGCG 845
Qy 723 GAAACGGGAAGCGGACCTGGAGTGCATGTCGGGTTTTCAACAAACCAATGCAAAATGCTGAAT 782
Db 844 GAAACGGGAAGCGGACCTGGAGTGCATGTCGGGTTTTCAACAAACCAATGCAAAATGCTGAAT 785
Qy 783 GAGGGCATCGTTCCCACTGCGATGCTGGTTGCGCAACGATCAGATGGCGTGGGCGCAATG 842
```

|||||
Db 784 GAGGGCATCGTTCCCACTGCGATGCTGGTGTGCCAACGATCAGATGCGCTGGGCGCAATG 725
Qy 843 CGCGCATTTACCGAGTCCGGGCTGGCGCTGGTGGCGGATATCTCGGTAGTGGATACGAC 902
Db 724 CGCGCATTTACCGAGTCCGGGCTGGCGCTGGTGGCGGATATCTCGGTAGTGGATACGAC 665
Qy 903 GATACCGAAGACAGCTCATGTTATATATCCCGCGTTAAACACCATCAAAACAGGATTTTCG 962
Db 664 GATACCGAAGACAGCTCATGTTATATATCCCGCGTTAAACACCATCAAAACAGGATTTTCG 605
Qy 963 CTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGAAG 1022
Db 604 CTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGAAG 545
Qy 1023 GGCATACAGCTGTTCCCGCTCTCACTGTGTGAAAGAAACACACCTCGCGCCCAATACG 1082
Db 544 GGCATACAGCTGTTCCCGCTCTCACTGTGTGAAAGAAACACACCTCGCGCCCAATACG 485
Qy 1083 CAAACCGGCTCTCCCGCGGTGGCGGATTAATGCACTGAGCTGGACGACAGGTTTCC 1142
Db 484 CAAACCGGCTCTCCCGCGGTGGCGGATTAATGCACTGAGCTGGACGACAGGTTTCC 425
Qy 1143 CGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTATTAGG 1202
Db 424 CGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGC-----GC 376
Qy 1203 ACAATTTCTCATGTTTGACAGCTTTATCATCGACTGCAAGCGGTGACCAATGCTTCTGGCGTC 1262
Db 375 GAAATTGATCTGTTTGACAGCTTTATCATCGACTGCAAGCGGTGACCAATGCTTCTGGCGTC 316
Qy 1263 AGGAGCCATCGGAAGCTGTGATGAGTGTGCGGTGTCAGGTCGTAATCACTGCATAATTCGTG 1322
Db 315 AGGAGCCATCGGAAGCTGTGATGAGTGTGCGGTGTCAGGTCGTAATCACTGCATAATTCGTG 256
Qy 1323 TCGCTCAAGGCGCACTCCGCTTCGATAATGTTTTTGGCGCGACATCAATACGGTTCT 1382
Db 255 TCGCTCAAGGCGCACTCCGCTTCGATAATGTTTTTGGCGCGACATCAATACGGTTCT 196
Qy 1383 GGCATAATTTCTGAATGAGCTGTTGACAAATTAATCAT--CGGCTCGTATAATGTGTGGAA 1441
Db 195 GGCATAATTTCTGAATGAGCTGTTGACAAATTAATCATCCGGCTGTATATGTGTGAA 136
Qy 1442 TTGTGAGCGGATAACAATTTTACACAGGAAACAGCC 1477
Db 135 TTGTGAGCGGATAACAATTTTACACAGGAAACAGAC 100

RESULT 5
US-09-640-882-2
; Sequence 2, Application US/09640882
; Patent No. 6720142
; GENERAL INFORMATION:
; APPLICANT: Hall, Barry G.
; TITLE OF INVENTION: METHOD OF DETERMINING EVOLUTIONARY POTENTIAL OF MUTANT
; TITLE OF INVENTION: RESISTANCE GENES AND USE THEREOF TO SCREEN FOR DRUG
; TITLE OF INVENTION: EFFICACY
; FILE REFERENCE: 176/60851
; CURRENT APPLICATION NUMBER: US/09/640,882
; CURRENT FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/149,813
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 5201
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid
; OTHER INFORMATION: pACSE2
US-09-640-882-2

Query Match 41.3%; Score 1362; DB 4; Length 5201;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1380; Conservative 0; Mismatches 10; Indels 2; Gaps 1;
Qy 3 GACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATATAGTCGCCGCCGAGAGAGT 62
Db 866 GACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATATAGTCGCCGCCGAGAGAGT 925
Qy 63 CAATTCAGGGTGGTGAATGTGAAACCCAGTAACTGATACGATGTGCGAGAGATATGCCGGT 122
Db 926 CAATTCAGGGTGGTGAATGTGAAACCCAGTAACTGATACGATGTGCGAGAGATATGCCGGT 985
Qy 123 GTCTCTTATCAGACGGTTTCCCGGTGTGTAACAGGCGCAGCGCTTCTCGGAAACG 182
Db 986 GTCTCTTATCAGACGGTTTCCCGGTGTGTAACAGGCGCAGCGCTTCTCGGAAACG 1045
Qy 183 CGGGAAAAGTGGAAAGCGGCGATGCGCGAGCTGAAATTAATATCCCAACCGGTGGCAAA 242
Db 1046 CGGGAAAAGTGGAAAGCGGCGATGCGCGAGCTGAAATTAATATCCCAACCGGTGGCAAA 1105
Qy 243 CAACTGGCGGCAAAACAGTCTGTTGATTCGCGGTGGCCACTCCAGTCTGGCCCTGCAC 302
Db 1106 CAACTGGCGGCAAAACAGTCTGTTGATTCGCGGTGGCCACTCCAGTCTGGCCCTGCAC 1165
Qy 303 GCGCGCTGCAAAATGTCGCGCGGATTAATATCTCGCGCGGATCAACTGGGTGCCAGCGTG 362
Db 1166 GCGCGCTGCAAAATGTCGCGCGGATTAATATCTCGCGCGGATCAACTGGGTGCCAGCGTG 1225
Qy 363 GTGGTGTGATGTAGAACGAGCGGCTGGAAGCTCTAAAGCGGCGGTGCACAACTTT 422
Db 1226 GTGGTGTGATGTAGAACGAGCGGCTGGAAGCTCTAAAGCGGCGGTGCACAACTTT 1285
Qy 423 CTGCGCAACCGTCACTGGGCTGATCAATTAATATCCGCTGGATGACAGGATGCCATT 482
Db 1286 CTGCGCAACCGTCACTGGGCTGATCAATTAATATCCGCTGGATGACAGGATGCCATT 1345
Qy 483 GCTGTGGAAGCTGCTGCACATAATGTTCCGGGCTTATTTCTTGATGTCTTGACAGACA 542
Db 1346 GCTGTGGAAGCTGCTGCACATAATGTTCCGGGCTTATTTCTTGATGTCTTGACAGACA 1405
Qy 543 CCCATCAACAGTATTATTTCTCCATGAAGAGTACGCGACTGGGCGTGAGGATCTG 602
Db 1406 CCCATCAACAGTATTATTTCTCCATGAAGAGTACGCGACTGGGCGTGAGGATCTG 1465
Qy 603 GTGCAATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTCGGCG 662
Db 1466 GTGCAATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTCGGCG 1525
Qy 663 GGTCTGCGTCTGGCTGGGTGGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
Db 1526 GGTCTGCGTCTGGCTGGGTGGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 1585
Qy 723 GAACGGGAAGCGACTGGAGTGCCATGTCGCGTTTCAACAAACCATGCAAAATGCTGAAT 782
Db 1586 GAACGGGAAGCGACTGGAGTGCCATGTCGCGTTTCAACAAACCATGCAAAATGCTGAAT 1645
Qy 783 GAGGCAATCGTTCCCACTGCGATGCTGTTGTCACACGATCAGATGCGGTGGGCGCAATG 842
Db 1646 GAGGCAATCGTTCCCACTGCGATGCTGTTGTCACACGATCAGATGCGGTGGGCGCAATG 1705
Qy 843 GCGGCATTACCGAGTCCGGGCTGCGCGTGGTGGGATATCTCGGTAGTGGGATACGAC 902
Db 1706 GCGGCATTACCGAGTCCGGGCTGCGCGTGGTGGGATATCTCGGTAGTGGGATACGAC 1765
Qy 903 GATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAACCATCAAAACAGGATTTTCG 962
Db 1766 GATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAACCATCAAAACAGGATTTTCG 1825
Qy 963 CTGCTGGGGCAAAACAGCGGTGGACCGCTTGCTGCAACTCTCTCAGGGCGCAGCGGTGAAG 1022
Db 1826 CTGCTGGGGCAAAACAGCGGTGGACCGCTTGCTGCAACTCTCTCAGGGCGCAGCGGTGAAG 1885
Qy 1023 GGCATACAGCTGTTGCGCGCTCTCACTGTGTGAAAGAAACCAACCCCTGGCGCCCAATACG 1082

```
Db 1886 GGAATTCAGCTGTGCGCGTCTCACTGGTGAAGAAACCAACCTGCGGCCAATACG 1945
Qy 1083 CAAACCGCCTCTCCCGCGCGTTGGCCGATTCAATTAATGACGCTGGCAGCAGGTTTCC 1142
Db 1946 CAAACCGCCTCTCCCGCGCGTTGGCCGATTCAATTAATGACGCTGGCAGCAGGTTTCC 2005
Qy 1143 CGACTGGAAGCGGCGAGTGAGCGCAACGAATTAATGTAGTTAGCTCACTATTAGGC 1202
Db 2006 CGACTGGAAGCGGCGAGTGAGCGCAACGAATTAATGTAGTTAG--CGCGAATTGATC 2063
Qy 1203 ACAATTTCTCATGTTTGAAGCTTATCATGCTGATCATGCTGACCGTGACCAATGCTTCTGGCGTC 1262
Db 2064 TGAATTTCTCATGTTTGAAGCTTATCATGCTGATCATGCTGACCGTGACCAATGCTTCTGGCGTC 2123
Qy 1263 AGGAGCCATCGGAAGCTGTGTGTATGGCTGTGAGTGTGCAAGTCTGTAATCACTGCAATAATTCGTG 1322
Db 2124 AGGAGCCATCGGAAGCTGTGTGTATGGCTGTGAGTGTGCAAGTCTGTAATCACTGCAATAATTCGTG 2183
Qy 1323 TGGCTCAAGGCGCACTCCGCTTCTGGAATAATGTTTTTTCGCGCGACATCAATAACGGTTCT 1382
Db 2184 TGGCTCAAGGCGCACTCCGCTTCTGGAATAATGTTTTTTCGCGCGACATCAATAACGGTTCT 2243
Qy 1383 GSCAAATATCT 1394
Db 2244 GSCAAATATCT 2255

RESULT 6
US-09-640-882-3
; Sequence 3, Application US/09640882
; Patent No. 6720142
; GENERAL INFORMATION:
; APPLICANT: Hall, Barry G.
; TITLE OF INVENTION: METHOD OF DETERMINING EVOLUTIONARY POTENTIAL OF MUTANT
; TITLE OF INVENTION: RESISTANCE GENES AND USE THEREOF TO SCREEN FOR DRUG
; TITLE OF INVENTION: EFFICACY
; FILE REFERENCE: 176/60851
; CURRENT APPLICATION NUMBER: US/09/640,882
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/149,813
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 5201
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid
US-09-640-882-3

Query Match 41.3%; Score 1362; DB 4; Length 5201;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1380; Conservative 0; Mismatches 10; Indels 2; Gaps 1;

Qy 3 GACACCATCGAATCGTGCMAAACCTTTTCGGGTATGGCATGATAGCCCGGAGAGAGT 62
Db 866 GACACCATCGAATCGTGCMAAACCTTTTCGGGTATGGCATGATAGCCCGGAGAGAGT 925
Qy 63 CAATTCAGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGAGAGTATGCCGT 122
Db 926 CAATTCAGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGAGAGTATGCCGT 985
Qy 123 GTCTCTTATCAGACCGTTTCCCGGTGGTGAACACAGCCAGCCAGCTTTCTTGGGAAACG 182
Db 986 GTCTCTTATCAGACCGTTTCCCGGTGGTGAACACAGCCAGCCAGCTTTCTTGGGAAACG 1045
Qy 183 CGGGAAGAGTGAAGCGGCGATGCGGAGCTGNAATTACATTCACACCGCGTGGCAAA 242
Db 1046 CGGGAAGAGTGAAGCGGCGATGCGGAGCTGNAATTACATTCACACCGCGTGGCAAA 1105
```

```
Qy 243 CAACCTGGCGGCAACAGTCTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGCAC 302
Db 1106 CAACCTGGCGGCAACAGTCTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGCAC 1165
Qy 303 GCGCGCTGCAAAATTTGTCGGCGGCGATTAAATCTCGCGCCGATCACTGGGTGCCAGCGTG 362
Db 1166 GCGCGCTGCAAAATTTGTCGGCGGCGATTAAATCTCGCGCCGATCACTGGGTGCCAGCGTG 1225
Qy 363 GTGCTGTCGATGTTAGAACGAGCGGCTGGAAGCCTGTAAAGCGGCGGTGCAAAATCTT 422
Db 1226 GTGCTGTCGATGTTAGAACGAGCGGCTGGAAGCCTGTAAAGCGGCGGTGCAAAATCTT 1285
Qy 423 CTCGCGCAACGCGTCAGTGGGCTGATCAATACTATCCGCTGGATGACCAAGGATGCCATT 482
Db 1286 CTCGCGCAACGCGTCAGTGGGCTGATCAATACTATCCGCTGGATGACCAAGGATGCCATT 1345
Qy 483 GCTGTGAAGCTGCTCCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACAGACA 542
Db 1346 GCTGTGAAGCTGCTCCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACAGACA 1405
Qy 543 CCCATCAACAGTATATTTTCTCCCATGAAGACGTTAGCGACTGGCGGTGGAGCATCTG 602
Db 1406 CCCATCAACAGTATATTTTCTCCCATGAAGACGTTAGCGACTGGCGGTGGAGCATCTG 1465
Qy 603 GTCGCAATTGGTCAACAGCAAAATCGCGCTGTAGCGGCGCCATTAACTTCTGTCGCGG 662
Db 1466 GTCGCAATTGGTCAACAGCAAAATCGCGCTGTAGCGGCGCCATTAACTTCTGTCGCGG 1525
Qy 663 GTCGCTGCTGGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAGCG 722
Db 1526 GTCGCTGCTGGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAGCG 1585
Qy 723 GAACGGGAAGCGCACTGGAGTGCATGTCGGGTTTTCACAAACCATGCAAAATGCTGAAT 782
Db 1586 GAACGGGAAGCGCACTGGAGTGCATGTCGGGTTTTCACAAACCATGCAAAATGCTGAAT 1645
Qy 783 GAGGGCATCGTCCCACTGCGATGCTGGTTGCCAAGCATCAGATGGCGCTGGCGCGCAATG 842
Db 1646 GAGGGCATCGTCCCACTGCGATGCTGGTTGCCAAGCATCAGATGGCGCTGGCGCGCAATG 1705
Qy 843 CGCGCATTAACGAGTCCGGGCTGGCGGTGGTGGCGGATATCTCGGTAGTGGGATACGAC 902
Db 1706 CGCGCATTAACGAGTCCGGGCTGGCGGTGGTGGCGGATATCTCGGTAGTGGGATACGAC 1765
Qy 903 GATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACCAACCATCAAAACAGGATTTTCG 962
Db 1766 GATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACCAACCATCAAAACAGGATTTTCG 1825
Qy 963 CTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGAAG 1022
Db 1826 CTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGAAG 1885
Qy 1023 GGAATCAGCTGTTGCCCGCTCACTGCTGTAAGAAAGAAACCAACCTTGGCGCCCAATACG 1082
Db 1886 GGAATCAGCTGTTGCCCGCTCACTGCTGTAAGAAAGAAACCAACCTTGGCGCCCAATACG 1945
Qy 1083 CAAACCGCCTCTCCCGCGCGTTGGCGGATTCAATTAATGAGCTGGCAGCAGAGTTTTC 1142
Db 1946 CAAACCGCCTCTCCCGCGCGTTGGCGGATTCAATTAATGAGCTGGCAGCAGAGTTTTC 2005
Qy 1143 CGACTGGAAGCGGCGAGTGAGCGCAACGAATTAATGTAGTTAGCTCACTATTAGGC 1202
Db 2006 CGACTGGAAGCGGCGAGTGAGCGCAACGAATTAATGTAGTTAG--CGCGAATTGATC 2063
Qy 1203 ACAATTTCTCATGTTTGAAGCTTATCATGCTGATCATGCTGACCGTGACCAATGCTTCTGGCGTC 1262
Db 2064 TGAATTTCTCATGTTTGAAGCTTATCATGCTGATCATGCTGACCGTGACCAATGCTTCTGGCGTC 2123
Qy 1263 AGGAGCCATCGGAAGCTGTGTATGGCTGTGAGTGTGCAAGTCTGTAATCACTGCAATAATTCGTG 1322
Db 2124 AGGAGCCATCGGAAGCTGTGTATGGCTGTGAGTGTGCAAGTCTGTAATCACTGCAATAATTCGTG 2183
Qy 1323 TCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCAATAACGGTTCT 1382
```

Db	2184	TCGCTCAAGCGCACTCCGTTCTGGATAATGTTTTTGGCCGCACATCATCAACGGTTCT	2243
Qy	1383	GGCAATATTCT 1394	
Db	2244	GGCAATATTCT 2255	
RESULT 7			
US-08-778-717-5			
: Sequence 5, Application US/08778717			
: Patent No. 602689			
: GENERAL INFORMATION:			
: APPLICANT: UENO, EIICHI			
: APPLICANT: NOBUYUKI, FUJII			
: APPLICANT: OKADA, MASAHISA			
: TITLE OF INVENTION: FUSED DNA SEQUENCE, FUSED PROTEIN			
: TITLE OF INVENTION: EXPRESSED FROM SAID FUSED DNA SEQUENCE AND METHOD FOR			
: TITLE OF INVENTION: EXPRESSING SAID DNA SEQUENCE			
: NUMBER OF INVENTIONS: 21			
: CORRESPONDENCE ADDRESS:			
: ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,			
: ADDRESSEE: P.C. S. JEFFERSON DAVIS HIGHWAY, SUITE 400			
: STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400			
: CITY: ARLINGTON			
: STATE: VA			
: COUNTRY: USA			
: ZIP: 22202			
: COMPUTER READABLE FORM:			
: MEDIUM TYPE: Floppy disk			
: COMPUTER: IBM PC compatible			
: OPERATING SYSTEM: PC-DOS/MS-DOS			
: SOFTWARE: PatentIn Release #1.0, Version #1.30			
: CURRENT APPLICATION DATA:			
: APPLICATION NUMBER: US/08/778,717			
: FILING DATE: 12-DEC-1996			
: CLASSIFICATION: 536			
: PRIOR APPLICATION DATA:			
: APPLICATION NUMBER: JP 352225/1995			
: FILING DATE: 28-DEC-1995			
: ATTORNEY/AGENT INFORMATION:			
: NAME: OBLON, NORMAN F.			
: REGISTRATION NUMBER: 24,618			
: REFERENCE/DOCKET NUMBER: 2084-031-0			
: TELECOMMUNICATION INFORMATION:			
: TELEPHONE: 703-413-3000			
: TELEFAX: 703-413-2220			
: INFORMATION FOR SEQ ID NO: 5:			
: SEQUENCE CHARACTERISTICS:			
: LENGTH: 4557 base pairs			
: TYPE: nucleic acid			
: STRANDEDNESS: double			
: TOPOLOGY: circular			
: MOLECULE TYPE: other nucleic acid			
: ORIGINAL SOURCE:			
: ORGANISM: E. COLI			
: STRAIN: BL21 (DE3)			
: PUBLICATION INFORMATION:			
: AUTHORS: NOBUYUKI FUJII ET AL,			
: TITLE: FUSED DNA SEQUENCE, FUSED PROTEIN EXPRESSED			
: TITLE: FROM SAID FUSED DNA SEQUENCE AND METHOD OF			
: TITLE: EXPRESSING SAID FUSED PROTEIN			
: RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 TO 4557			
US-08-778-717-5			
Query Match 36.5%; Score 1204; DB 4; Length 4557;			
Best Local Similarity 100.0%; Pred.No. 0;			
Matches 1204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCCCGCGAAGAGA	60
Db	2563	CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCCCGCGAAGAGA	2622

|||||
Db 3703 CCGGACTGAAAGCGGCGAGTGGCGCAACGAATTAATGTAGTACTCACTATTAG 3762
|||
Qy 1201 GCAC 1204
|||
Db 3763 GCAC 3766
|||
RESULT 8
US-08-148-675A-2
; Sequence 2, Application US/08148675A
; Patent No. 5506121
; GENERAL INFORMATION:
; APPLICANT: Skerra, Arne; Schmidt, Thomas
; TITLE OF INVENTION: FUSION PEPTIDES WITH BINDING ACTIVITY FOR
; TITLE OF INVENTION: STREPTAVIDIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESS: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/148,675A
; FILING DATE: 3-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 42 37 113.9
; FILING DATE: 03-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsai, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: HUBR 1041
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3832 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-148-675A-2
Query Match 36.4%; Score 1200.8; DB 1; Length 3832;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1202; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGCAAAACCTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 3 CCGACACCATCGAATGGTGCAAAACCTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA 62
Qy 61 GTCAATTCAGGGTGGTGAATGTGAACAGTAACGTTTATACGATGTCGACAGATATGCCG 120
Db 63 GTCAATTCAGGGTGGTGAATGTGAACAGTAACGTTTATACGATGTCGACAGATATGCCG 122
Qy 121 GTGTCCTTATCAGACCGTTTCGCGGTGTGAACAGCGACCGACCGTTTCTCGCAAAA 180
Db 123 GTGTCCTTATCAGACCGTTTCGCGGTGTGAACAGCGACCGTTTCTCGCAAAA 182
Qy 181 CGCGGGAAGTGAAGCGCGCATGCGGAGCTGAATTCATTCCTCAACCGGTGGCAC 240
Db 183 CGCGGGAAGTGAAGCGCGCATGCGGAGCTGAATTCATTCCTCAACCGGTGGCAC 242
Qy 241 AACACTGGCGGCAACACAGTCGTTGTGATTTGGCGTTGCCACTCCAGTCTGGCCCTGC 300
|||||

Db 243 AACAACTGGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 302
Qy 301 ACGCCCGTTCGCAAAATTTGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db 303 ACGCCCGTTCGCAAAATTTGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 362
Qy 361 TGGTGGTGTGATGTTAGAACGAAAGCGCGTTCGAAGCCTCTGAAGCGCGGTGCACAATC 420
Db 363 TGGTGGTGTGATGTTAGAACGAAAGCGCGTTCGAAGCCTCTGAAGCGCGGTGCACAATC 422
Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACACAGATGCCA 480
Db 423 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACACAGATGCCA 482
Qy 481 TTGCTGTGGAGCTGCGCTGCACTAATGTTCCGGGTTATTTCTTGATGTCCTTGACACAGA 540
Db 483 TTGCTGTGGAGCTGCGCTGCACTAATGTTCCGGGTTATTTCTTGATGTCCTTGACACAGA 542
Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Db 543 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 602
Qy 601 TGGTCGCAATGGGTCACACAGAAATCGCGCTGTAGCGGCCCCATTAAGTTCTGTCTCGG 660
Db 603 TGGTCGCAATGGGTCATCAGCAAAATCGCGCTGTAGCGGCCCCATTAAGTTCTGTCTCGG 662
Qy 661 CGGCTCTCGCTGCGCTGGCTGGCGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 663 CGGCTCTCGCTGCGCTGGCTGGCGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 722
Qy 721 CGGAACGGGAAGGCGACTGGAGTCCCATGTCGGGTTTCAACAAACCATCAAAATGCTGA 780
Db 723 CGGAACGGGAAGGCGACTGGAGTCCCATGTCGGGTTTCAACAAACCATCAAAATGCTGA 782
Qy 781 ATGAGGCGATCGTTTCCCATGTCGCGATGCTGTTGCCAAACGATCAGATGGCGCTGGCGCAA 840
Db 783 ATGAGGCGATCGTTTCCCATGTCGCGATGCTGTTGCCAAACGATCAGATGGCGCTGGCGCAA 842
Qy 841 TGGCGCGCATTCAGCGAGTCGGGCTGGGCTGGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 843 TGGCGCGCATTCAGCGAGTCGGGCTGGGCTGGGTTGGTGGGATATCTCGGTAGTGGGATACG 902
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAACCATCAAAACAGGATTTTC 960
Db 903 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAACCATCAAAACAGGATTTTC 962
Qy 961 GCCTGCTGGGGCAAAACAGGCTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGATGA 1020
Db 963 GCCTGCTGGGGCAAAACAGGCTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGATGA 1022
Qy 1021 AGGGCAATCAGCTGTTGCGCGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGGCCCAATA 1080
Db 1023 AGGGCAATCAGCTGTTGCGCGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGGCCCAATA 1082
Qy 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGACAGCTGGCAACGACAGGTTT 1140
Db 1083 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGACAGCTGGCAACGACAGGTTT 1142
Qy 1141 CCGGACTGGAAGCGGGCAGTGAAGCGCAACGCAATTAATGTGAGTTAGTCTCACTATTAG 1200
Db 1143 CCGGACTGGAAGCGGGCAGTGAAGCGCAACGCAATTAATGTGAGTTAGTCTCACTATTAG 1202
Qy 1201 GCAC 1204
Db 1203 GCAC 1206

RESULT 9
US-08-487-283A-18/c
; Sequence 18, Application US/08487283A
; Patent No. 6355245
; GENERAL INFORMATION:
; APPLICANT: Evans, Mark J.

APPLICANT: Matis, Louis A.
APPLICANT: Mueller, Eileen Elliott
APPLICANT: Nye, Steven H.
APPLICANT: Rollins, Scott
APPLICANT: Rother, Russell P.
APPLICANT: Springhorn, Jeremy P.
APPLICANT: Suinto, Stephen P.
APPLICANT: Thomas, Thomas C.
APPLICANT: Wilkins, James A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT
OF INFLAMMATORY DISEASES
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSER: Seth A. Fidel
STREET: 25 Science Park (Alexion)
CITY: New Haven
STATE: Connecticut
COUNTRY: USA
ZIP: 06511
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.4Mb storage
COMPUTER: Macintosh Cetrus 610
OPERATING SYSTEM: System 7
SOFTWARE: WordPerfect 3.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,283A
FILING DATE: June 7, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/236,208
FILING DATE: 02-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Seth A. Fidel.
REGISTRATION NUMBER: 38,449
REFERENCE/DOCKET NUMBER: ALX-152.1 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203)776-1790
TELEFAX: (203)772-3655
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 5248 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Circular
MOLECULE TYPE: Other nucleic acid
DESCRIPTION: pET Trc SOS/NI
DESCRIPTION: prokaryotic expression vector
US-08-487-283A-18

Query Match 36.3%; Score 1199.2; DB 3; Length 5248;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CGGACACCATCAATGGTCAAAACCTTTCCGGGTATGGCATGATAGCCCGGAAGAGA 60
DB 4810 CGGACACCATCAATGGCGCAAAACCTTTCCGGGTATGGCATGATAGCCCGGAAGAGA 4751
QY 61 GTCAATTACAGGTGGTGAATGTGAACACAGTAAACGTTATACGATGTCGAGAGTATGCCG 120
DB 4750 GTCAATTACAGGTGGTGAATGTGAACACAGTAAACGTTATACGATGTCGAGAGTATGCCG 4691
QY 121 GTGTCTCTTATCAGACCGTTTCCGGGTGTGAACACGAGCCGACGCTTTCTGGGAAA 180
DB 4690 GTGTCTCTTATCAGACCGTTTCCGGGTGTGAACACGAGCCGACGCTTTCTGGGAAA 4631
QY 181 CGCGGAAAAGTGAAGCGGCGATGGCGAGCTGAATTACATTCACACCGCGTGGCAC 240
DB 4630 CGCGGAAAAGTGAAGCGGCGATGGCGAGCTGAATTACATTCACACCGCGTGGCAC 4571
QY 241 AACAACTGGCGGCAACAGTGTGCTGTAATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
DB 4570 AACAACTGGCGGCAACAGTGTGCTGTAATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 4511
QY 301 ACGCGCCGTGCGAAATTTGTCGGCGGCGAATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 360

Db 4510 ACGCGCCGTGCGAAATTTGTCGGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 4451
QY 361 TGGTGGTGTGATGGTGAACGAAGCGCGTGAAGCGCTTAAAGCGCGGTGCAACAATC 420
Db 4450 TGGTGGTGTGATGGTGAACGAAGCGCGTGAAGCGCTTAAAGCGCGGTGCAACAATC 4391
QY 421 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATCACCAGGATGCCA 480
Db 4390 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATCACCAGGATGCCA 4331
QY 481 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCCGGCGTATTCTTCTGATGCTCTGACACAGA 540
Db 4330 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCCGGCGTATTCTTCTGATGCTCTGACACAGA 4271
QY 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACCGGACTGGGCGGTGGAGCATC 600
Db 4270 CACCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACCGGACTGGGCGGTGGAGCATC 4211
QY 601 TGGTGGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db 4210 TGGTGGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 4151
QY 661 CGCGTCTCGCTCTGGCTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4150 CGCGTCTCGCTCTGGCTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 4091
QY 721 CGGAACGGGAAGGGGACTGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db 4090 CGGAACGGGAAGGGGACTGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 4031
QY 781 ATGAGGGCATGTTCCCACTGGGATGCTGGTTCGCAACGATCAGATGGGCTGGGCGGCA 840
Db 4030 ATGAGGGCATGTTCCCACTGGGATGCTGGTTCGCAACGATCAGATGGGCTGGGCGGCA 3971
QY 841 TGGCGCCATTTACCGAGTCCGGGCTGGCGTGGTGGCGGATATCTCGGTAGTGGGATAGC 900
Db 3970 TGGCGCCATTTACCGAGTCCGGGCTGGCGTGGTGGCGGATATCTCGGTAGTGGGATAGC 3911
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACCAACCATCAAAAGGATTTTC 960
Db 3910 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACCAACCATCAAAAGGATTTTC 3851
QY 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGTGCAAACTCTCTCAGGCGCCAGGCGGTGA 1020
Db 3850 GCCTGCTGGGCAACACAGCGTGGACCGCTTGTGCAAACTCTCTCAGGCGCCAGGCGGTGA 3791
QY 1021 AGGGCAATCAGCTGTTGCCCGCTCTCACCTGGTGAAGAAACACACCTGGCGCCCAATA 1080
Db 3790 AGGGCAATCAGCTGTTGCCCGCTCTCACCTGGTGAAGAAACACACCTGGCGCCCAATA 3731
QY 1081 CGCAACCGCCCTCTCCCGCGCGTTGGCGGATTCATTAATGCAGCTGGCACGACAGGTTT 1140
Db 3730 CGCAACCGCCCTCTCCCGCGCGTTGGCGGATTCATTAATGCAGCTGGCACGACAGGTTT 3671
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTGTAGTCTCATTTAG 1200
Db 3670 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTGTAGTCTCATTTAG 3611
QY 1201 GCAC 1204
Db 3610 GCAC 3607
RESULT 10
PCT-US96-05611A-21/c
; Sequence 21, Application PC/TUS9605611A
; GENERAL INFORMATION:
; APPLICANT: Mueller, John P.
; APPLICANT: Leonardo, Michael J.
; APPLICANT: McFarland, Henry P.
; APPLICANT: Matis, Louis A.
; APPLICANT: Mueller, Eileen Elliott

; APPLICANT: Nye, Steven H.
 ; APPLICANT: Pelfrey, Clara M.
 ; APPLICANT: Squinto, Stephen P.
 ; APPLICANT: Wilkins, James A.
 ; TITLE OF INVENTION: Modified Myelin Protein Molecules
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Maurice M. Klee
 ; STREET: 1951 Burr Street
 ; CITY: Fairfield
 ; STATE: Connecticut
 ; COUNTRY: USA
 ; ZIP: 06430
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 inch, 0.8 Mb storage
 ; COMPUTER: Macintosh Centris 610
 ; OPERATING SYSTEM: System 7
 ; SOFTWARE: Microsoft Word 6.0.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US96/05611A
 ; FILING DATE: 02-MAY-1995
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/431,644
 ; FILING DATE: May 2, 1995
 ; APPLICATION NUMBER: 08/431,648
 ; FILING DATE: May 2, 1995
 ; APPLICATION NUMBER: 08/482,114
 ; FILING DATE: June 7, 1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Klee, Maurice M.
 ; REGISTRATION NUMBER: 30,399
 ; REFERENCE/DOCKET NUMBER: ALX-129
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (203) 255 1400
 ; TELEFAX: (203) 254 1101
 ; INFORMATION FOR SEQ ID NO: 21:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 5248 base pairs
 ; TYPE: Nucleic Acid
 ; STRANDEDNESS: Double
 ; TOPOLOGY: Circular
 ; MOLECULE TYPE: Other nucleic acid
 ; DESCRIPTION: pet trc SOS/NI
 ; DESCRIPTION: prokaryotic expression vector
 ; PCT-US96-05611A-21

Query Match 36.3%; Score 1199.2; DB 5; Length 5248;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 1 CCGACACCATCGAATGGTGCAGGAGCTTTCGCGGTATGGCATGATAGCGCCGGAAGAGA 60
 DB 4810 CGGACACCATCGAATGGCGCAAAACCTTTCGCGGTATGGCATGATAGCGCCGGAAGAGA 4751
 QY 61 GTCAATTCAGGGTGGTGAATGTGAAACCAAGTAACGTTATACGATGTGCGAGAGTATCCCG 120
 DB 4750 GTCAATTCAGGGTGGTGAATGTGAAACCAAGTAACGTTATACGATGTGCGAGAGTATCCCG 4691
 QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGTAACGAGGCGCAGCCAGTTCCTCGGAAA 180
 DB 4690 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGTAACGAGGCGCAGCCAGTTCCTCGGAAA 4631
 QY 181 CGCGGGAAAAAGTGGAGCGCGCATGCGGAGCTGAATTACATTCGCCAACCGCGTGGCAC 240
 DB 4630 CGCGGGAAAAAGTGGAGCGCGCATGCGGAGCTGAATTACATTCGCCAACCGCGTGGCAC 4571
 QY 241 AACCACTGGCGGCAACAGTCGTTGCTGATGGCGTTGCCACTCCAGTCTGCCCTCGC 300
 DB 4570 AACCACTGGCGGCAACAGTCGTTGCTGATGGCGTTGCCACTCCAGTCTGCCCTCGC 4511
 QY 301 ACGCGCGTGCAGAAATGTTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 360

DB 4510 ACGCGCGTGCAGAAATGTTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGC 4451
 QY 361 TGGTGGTGTGATCGTAGAAGAGCGCGTGCAGAGCTGTAAAGCGCGCGGTGCACAAATC 420
 DB 4450 TGGTGGTGTGATCGTAGAAGAGAGCGCGTGCAGAGCTGTAAAGCGCGCGGTGCACAAATC 4391
 QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAACATATCCGCTCGATGACCGAGATGCCA 480
 DB 4390 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAACATATCCGCTCGATGACCGAGATGCCA 4331
 QY 481 TTGCTGTGGAAGCTGCCTGCACATAATGTTCCGGGCTTAATTTCTTGATGTCTCTGACCAGA 540
 DB 4330 TTGCTGTGGAAGCTGCCTGCACATAATGTTCCGGGCTTAATTTCTTGATGTCTCTGACCAGA 4271
 QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 600
 DB 4270 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 4211
 QY 601 TGGTCGCATTTGGGTACAGCAAAATCGCGTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
 DB 4210 TGGTCGCATTTGGGTACAGCAAAATCGCGTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 4151
 QY 661 CGGCTCTGCGTCTGGGTGGCGATAAATATCTCACTCGCAATCAAAATTCAGCGCATAG 720
 DB 4150 CGGCTCTGCGTCTGGGTGGCGATAAATATCTCACTCGCAATCAAAATTCAGCGCATAG 4091
 QY 721 CGGAACGGGAAGCGACTGGAGTGCCATGTCGGTTCCTCAACCAACCATGCAATGCTGA 780
 DB 4090 CGGAACGGGAAGCGACTGGAGTGCCATGTCGGTTCCTCAACCAACCATGCAATGCTGA 4031
 QY 781 ATGAGGCGCATGTTCCCACTGCGATGCTGTTGCAACAGATCAGATGGCGCTGGGCGCAA 840
 DB 4030 ATGAGGCGCATGTTCCCACTGCGATGCTGTTGCAACAGATCAGATGGCGCTGGGCGCAA 3971
 QY 841 TGGCGCGCATTAACGAGTCCGGCTGCGCTTGGTGGGATATCTCGGTAGTGGGATACG 900
 DB 3970 TGGCGCGCATTAACGAGTCCGGCTGCGCTTGGTGGGATATCTCGGTAGTGGGATACG 3911
 QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTCACCAACCATCAACAGATTTTC 960
 DB 3910 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTCACCAACCATCAACAGATTTTC 3851
 QY 961 GCCTGCTGGGGCAAAACCGAGTCCGCTTGTGCAAACTCTCTCAGGGCCAGCGGTGA 1020
 DB 3850 GCCTGCTGGGGCAAAACCGAGTCCGCTTGTGCAAACTCTCTCAGGGCCAGCGGTGA 3791
 QY 1021 AGGCGAATCAGCTGTTGCGCGTCTCACTGGTGAAGAAAAACCAACCCCTGGCGCCCAATA 1080
 DB 3790 AGGCGAATCAGCTGTTGCGCGTCTCACTGGTGAAGAAAAACCAACCCCTGGCGCCCAATA 3731
 QY 1081 CGCAACCGCGCTCTCCCGCGGTTGCGCGATTCATTATGACAGCTGGCAGCAGCGTTT 1140
 DB 3730 CGCAACCGCGCTCTCCCGCGGTTGCGCGATTCATTATGACAGCTGGCAGCAGCGTTT 3671
 QY 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACCGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
 DB 3670 CCCGACTGGAAGCGGCGAGTGAAGCGCAACCGCAATTAATGTGAGTTAGTCACTCATTTAG 3611
 QY 1201 GCAC 1204
 DB 3610 GCAC 3607

RESULT 11
 US-10-263-103-35/c
 ; Sequence 35, Application US/10263103
 ; Patent No. 6821723
 ; GENERAL INFORMATION:
 ; APPLICANT: AVENTIS PASTEUR
 ; APPLICANT: Chevalier, Michel
 ; APPLICANT: El Habib, Raphaelle
 ; APPLICANT: Krell, Tino
 ; APPLICANT: Sodoyer, Regis

; TITLE OF INVENTION: gp41 antigen
; FILE REFERENCE: 01-1692-A
; CURRENT APPLICATION NUMBER: US/10/263,103
; CURRENT FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 5312
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: plasmid pBT-cer
US-10-263-103-35

Query Match 36.3%; Score 1199.2; DB 4; Length 5312;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	1	CCGACACCATCGAATGGTGC	AAACCTTTCGGGTATGGCATGATAGCGCCGGAAGAGA	60
DB	4622	CGGACACCATCGAATGGCGCA	AAACCTTTCGGCGGTATGGCATGATAGCGCCGGAAGAGA	4563
QY	61	GTCAATTCAAGGTGGTGAAT	GTGAACACGAGTAACGTTATACGATGTCGCAGAGTATGCCG	120
DB	4562	GTCAATTCAAGGTGGTGAAT	GTGAACACGAGTAACGTTATACGATGTCGCAGAGTATGCCG	4503
QY	121	GTGTCCTTTATCAGACCGG	TTTCCCGGTGGTGAACACGAGCCAGCCACGTTTCTGCGAAAA	180
DB	4502	GTGTCCTTTATCAGACCGG	TTTCCCGGTGGTGAACACGAGCCAGCCACGTTTCTGCGAAAA	4443
QY	181	CGCGGGAAGAGTGAAGCGG	CGATGGCGGAGCTGAATTACATTCGCCAACCGGTGGCAC	240
DB	4442	CGCGGGAAGAGTGAAGCGG	CGATGGCGGAGCTGAATTACATTCGCCAACCGGTGGCAC	4383
QY	241	AACAACTGCGGGCAAAAG	TCGTTGCTGATTGGGGTTGCCACCTCCAGTCTGGGCCCTGC	300
DB	4382	AACAACTGCGGGCAAAAG	TCGTTGCTGATTGGGGTTGCCACCTCCAGTCTGGGCCCTGC	4323
QY	301	AGCGCCGTCGCAAAATGT	CGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG	360
DB	4322	AGCGCCGTCGCAAAATGT	CGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG	4263
QY	361	TGGTGGTTCGATGTAGNA	CGAAGCGGGCTCGAAGCCTGTAAAGCGGGGTGCAATC	420
DB	4262	TGGTGGTTCGATGTAGNA	CGAAGCGGGCTCGAAGCCTGTAAAGCGGGGTGCAATC	4203
QY	421	TTCTCGCGCAACGCTCAG	TGGGCTGATCAATTAATCTCCGCTGGATGACCAAGGATGCA	480
DB	4202	TTCTCGCGCAACGCTCAG	TGGGCTGATCAATTAATCTCCGCTGGATGACCAAGGATGCA	4143
QY	481	TTGCTGTGGAAGCTGCCT	GCACTAATGTTCCGGCGTTATTCTTGATGTCTCTGACCCAGA	540
DB	4142	TTGCTGTGGAAGCTGCCT	GCACTAATGTTCCGGCGTTATTCTTGATGTCTCTGACCCAGA	4083
QY	541	CACCCATCAAGTATTATTT	CTCCATGAAGACGGTACGCGATCGGCGGTGGAGCATC	600
DB	4082	CACCCATCAAGTATTATTT	CTCCATGAAGACGGTACGCGATCGGCGGTGGAGCATC	4023
QY	601	TGGTGCATTTGGTCCAGC	AAATCGCGCTGTAGCGGGCCCATTAAGTTCTGCTCGG	660
DB	4022	TGGTGCATTTGGTCCAGC	AAATCGCGCTGTAGCGGGCCCATTAAGTTCTGCTCGG	3963
QY	661	CCCGTCTGCGTCTGGCTG	CGCTAAATATCTCACTCGCAATCAAAATTCACGCGATAG	720
DB	3962	CCCGTCTGCGTCTGGCTG	CGCTAAATATCTCACTCGCAATCAAAATTCACGCGATAG	3903
QY	721	CGGAAACGGGAACGCGAT	GCGATGTCGTTTCAACAAACCAATGCAAAATGCTGA	780
DB	3902	CGGAAACGGGAACGCGAT	GCGATGTCGTTTCAACAAACCAATGCAAAATGCTGA	3843
QY	781	ATGAGGGGATCGTCCCA	CTCGGATGTTGCCAAGATGCGGTGGCGGCA	840
DB	3842	ATGAGGGGATCGTCCCA	CTCGGATGTTGCCAAGATGCGGTGGCGGCA	3783

QY	841	TGCGGCGCATTTACGAGT	CCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGATACG	900
DB	3782	TGCGGCGCATTTACGAGT	CCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGATACG	3723
QY	901	ACGATACCGAAGACAGCT	CAATGTTATATCCCGCGTAAACACCATCAAAACAGGATTTTC	960
DB	3722	ACGATACCGAAGACAGCT	CAATGTTATATCCCGCGTAAACACCATCAAAACAGGATTTTC	3663
QY	961	GCCTGCTGGGCAAAACCA	CGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020
DB	3662	GCCTGCTGGGCAAAACCA	CGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA	3603
QY	1021	AGGGCAATCAGCTGTTGC	CCCTCTCACTGGTGAAGAAAAACCACTTGGGCGCCCAATA	1080
DB	3602	AGGGCAATCAGCTGTTGC	CCCTCTCACTGGTGAAGAAAAACCACTTGGGCGCCCAATA	3543
QY	1081	CGCAAAACCGCTCTCCC	CGCGTGGCGGTTGGCGGATTCATTAAATGAGTGAAGTTAGTCACTCATTAG	1140
DB	3542	CGCAAAACCGCTCTCCC	CGCGTGGCGGTTGGCGGATTCATTAAATGAGTGAAGTTAGTCACTCATTAG	3483
QY	1141	CCCGACTGAAAAGCGGG	CGAGTGAGCGCAACGCAATTAATGTAGTGAAGTTAGTCACTCATTAG	1200
DB	3482	CCCGACTGAAAAGCGGG	CGAGTGAGCGCAACGCAATTAATGTAGTGAAGTTAGTCACTCATTAG	3423
QY	1201	GCAC	1204	
DB	3422	GCAC	3419	

RESULT 12

US-08-929-967-1/c
; Sequence 1, Application US/08929967
; Patent No. 5891637
; GENERAL INFORMATION:
; APPLICANT: Ruppert, Siegfried J.W.
; TITLE OF INVENTION: Construction of Full-length cDNA Libraries
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/929,967
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1035R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5443 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-929-967-1

Query Match 36.3%; Score 1199.2; DB 2; Length 5443;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAGAACCTTTCCGGGTATGGCATGTAGCCCGGAAGAGA 60
Db CGGACACCATCGAATGGCGCAAAACCTTTCCGGGTATGGCATGTAGCCCGGAAGAGA 4751
QY 61 GTCAATTCAGGGTGGTGAATGTGAACACAGTAACTGATATACGATGCGAGAGTATGCCG 120
Db GTCAATTCAGGGTGGTGAATGTGAACACAGTAACTGATATACGATGCGAGAGTATGCCG 4691
QY 121 GTGTCTCTTATCAGACCGCTTTCCGGGTGGTGAACCCAGGCCAGCCAGCTTTCTGCGAAA 180
Db GTGTCTCTTATCAGACCGCTTTCCGGGTGGTGAACCCAGGCCAGCCAGCTTTCTGCGAAA 4631
QY 181 CCGGGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTTACATTCCTCAACCCGCTGGCAC 240
Db CCGGGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTTACATTCCTCAACCCGCTGGCAC 4571
QY 241 AACCACTGGCGGCAACAGTGTGCTGATTTGGGTTGGCCACCTCCAGTCTGGCCCTGC 300
Db AACCACTGGCGGCAACAGTGTGCTGATTTGGGTTGGCCACCTCCAGTCTGGCCCTGC 4511
QY 301 ACGCGCGTGCAGAAATTTGTCGGGCGATTAATCTCGGCGCGATCAACTGGGTGCCAGG 360
Db ACGCGCGTGCAGAAATTTGTCGGGCGATTAATCTCGGCGCGATCAACTGGGTGCCAGG 4451
QY 361 TGGTGGTTCGATGTAGAACGAGCGGCTGAAAGCTGTAAAGCGCGGTGCACAAATC 420
Db TGGTGGTTCGATGTAGAACGAGCGGCTGAAAGCTGTAAAGCGCGGTGCACAAATC 4391
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGTCGATGCTCTGACACAGA 480
Db TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGTCGATGCTCTGACACAGA 4331
QY 481 TTGCTGTGGAAGCTGCCCTGCAATAATGTTCGGGCGTATTTCTTGTGCTCTGACACAGA 540
Db TTGCTGTGGAAGCTGCCCTGCAATAATGTTCGGGCGTATTTCTTGTGCTCTGACACAGA 4271
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAGACGGTACCGCATCGGCGGTGAGCATC 600
Db CACCCATCAACAGTATTAATTTCTCCCATGAGACGGTACCGCATCGGCGGTGAGCATC 4211
QY 601 TGGTGCATTTGGTTCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTCTCTCGG 660
Db TGGTGCATTTGGTTCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTCTCTCGG 4151
QY 661 CGGCTCTCGCTCGCTGGCATTAATCTCACTCGCAATCAAAATCAGCCGATAG 720
Db CGGCTCTCGCTCGCTGGCATTAATCTCACTCGCAATCAAAATCAGCCGATAG 4091
QY 721 CCGAACCGGAAGGCGACTGGAGTGCCATGTCCGGTTTTCACCAACCATGCAATGCTGA 780
Db CCGAACCGGAAGGCGACTGGAGTGCCATGTCCGGTTTTCACCAACCATGCAATGCTGA 4091
QY 841 ATGAGGGCATGTTCCCATCTCGATGCTGGTTGCAACGATCAGATGGCGCTGGCGCAA 840
Db ATGAGGGCATGTTCCCATCTCGATGCTGGTTGCAACGATCAGATGGCGCTGGCGCAA 3971
QY 841 TGGCGGCATTTACCGAGTCCGGGCTGGCTGGTGGGATATCTCGGTAGTGGATACG 900
Db TGGCGGCATTTACCGAGTCCGGGCTGGCTGGTGGGATATCTCGGTAGTGGATACG 3971
QY 901 ACGATACCGAAGA CAGCTCATGTTATATCCCGCGTTTAAACCATCAACAGATTTTC 960
Db ACGATACCGAAGA CAGCTCATGTTATATCCCGCGTTTAAACCATCAACAGATTTTC 3951
QY 961 GCCTGCTGGGGCAAC CAGCGTGAACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db GCCTGCTGGGGCAAC CAGCGTGAACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 3791
QY 1021 AGGGCAATCAGCTGTTGCGCGTCTCACTGGTGAAGAAACACCGCTGGCGCCCAATA 1080
Db AGGGCAATCAGCTGTTGCGCGTCTCACTGGTGAAGAAACACCGCTGGCGCCCAATA 3731
QY 1081 CGCAACCGGCTCTCCCGCGCGCTTGGCGGATTCATTAATGCACTGGCACGACGAGTTT 1140

Db 3730 CGCAACCGGCTCTCCCGCGCTTGGCGGATTCATTAATGCACTGGCACGACAGGTTT 3671
QY 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTCTAGCTCATTAG 1200
Db 3670 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTCTAGCTCATTAG 3611
QY 1201 GCAC 1204
Db 3610 GCAC 3607
RESULT 13
US-09-702-705-785/c
; Sequence 785, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: DIAGNOSIS OF LUNG CANCER
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-785

Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGCAGAACCTTTCCGGGTATGGCATGTAGCCCGGAAGAGA 60
Db 4677 CCGACACCATCGAATGGCGCAAAACCTTTCCGGGTATGGCATGTAGCCCGGAAGAGA 4618
QY 61 GTCAATTCAGGGTGGTGAATGTGAACCCAGTAACTGATATACGATGTCGAGAGTATGCCG 120
Db 4617 GTCAATTCAGGGTGGTGAATGTGAACCCAGTAACTGATATACGATGTCGAGAGTATGCCG 4558
QY 121 GTGTCTCTTATCAGACCGCTTTCCCGGTGGTGAACCCAGGCCAGCCAGCTTTCTGCGAAA 180
Db 4557 GTGTCTCTTATCAGACCGCTTTCCCGGTGGTGAACCCAGGCCAGCCAGCTTTCTGCGAAA 4498
QY 181 CCGGGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTTACATTCCTCAACCCGCTGGCAC 240
Db 4497 CCGGGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTTACATTCCTCAACCCGCTGGCAC 4438
QY 241 AACCACTGGCGGCAAC CAGTGTGCTGATTTGGGTTGGCCACCTCCAGTCTGGCCCTGC 300
Db 4437 AACCACTGGCGGCAAC CAGTGTGCTGATTTGGGTTGGCCACCTCCAGTCTGGCCCTGC 4378
QY 301 ACGCGCGTGCAGAAATTTGTCGGGCGATTAATCTCGGCGCGATCAACTGGGTGCCAGG 360
Db 4377 ACGCGCGTGCAGAAATTTGTCGGGCGATTAATCTCGGCGCGATCAACTGGGTGCCAGG 4318
QY 361 TGGTGGTTCGATGTAGAACGAGCGGCTGAAAGCTGTAAAGCGCGGTGCACAAATC 420
Db 4317 TGGTGGTTCGATGTAGAACGAGCGGCTGAAAGCTGTAAAGCGCGGTGCACAAATC 4258
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTCATTAATGCACTGGCATGCAACGAGTTT 480

```
Db 4257 TTCTCGGCAACGCGTCAGTGGCTGATCAATTAACATATCCGCTGGATGACACGAGTGC 4198
QY 481 TTGCTGTGGAAGCTCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCCTGACGAGA 540
Db 4197 TTGCTGTGGAAGCTCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCCTGACGAGA 4138
QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGGCACTGGCGCTGGAGCATC 600
Db 4137 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGGCACTGGCGCTGGAGCATC 4078
QY 601 TGGTCGCATTTGGTCAACAGCAAAATCGCGCTGTAGCGGCGCCATTAAGTTCTGTCTCGG 660
Db 4077 TGGTCGCATTTGGTCAACAGCAAAATCGCGCTGTAGCGGCGCCATTAAGTTCTGTCTCGG 4018
QY 661 CGCGTCTGCGTCTGCGTGGCTGGCAATAATATCTCACTGCGCAATCAAAATTCAGCCGATAG 720
Db 4017 CGCGTCTGCGTCTGCGTGGCTGGCAATAATATCTCACTGCGCAATCAAAATTCAGCCGATAG 3958
QY 721 CGGAACGGGAAGGCACTGGAGTGGCATGTCGGTTCCTCAAAACCATGCAAAATGCTGA 780
Db 3957 CGGAACGGGAAGGCACTGGAGTGGCATGTCGGTTCCTCAAAACCATGCAAAATGCTGA 3898
QY 781 ATGAGGGCATGTTCCCACTCGGATGCTGGTTGCCAAGCATGAGATGGCGCTGGCGCAA 840
Db 3897 ATGAGGGCATGTTCCCACTCGGATGCTGGTTGCCAAGCATGAGATGGCGCTGGCGCAA 3838
QY 841 TCGCGGCCATTACCGAGTCCGGGCTCGCGTGGTGCGGATATCTCGGTAGTGGGATAG 900
Db 3837 TCGCGGCCATTACCGAGTCCGGGCTCGCGTGGTGCGGATATCTCGGTAGTGGGATAG 3778
QY 901 AGGATACGGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAATCAAAACAGGATTTTC 960
Db 3777 AGGATACGGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAATCAAAACAGGATTTTC 3718
QY 961 GCCTGCTGGGGCAACACGCGTGGACCGCTTCTGCACTCTCTCAGGCGCACGCGGTGA 1020
Db 3717 GCCTGCTGGGGCAACACGCGTGGACCGCTTCTGCACTCTCTCAGGCGCACGCGGTGA 3658
QY 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGCTGAAAGAAACCAACCTGGCGCCCAATA 1080
Db 3657 AGGGCAATCAGCTGTTGCCCGTCTCACTGCTGAAAGAAACCAACCTGGCGCCCAATA 3598
QY 1081 CGCAACCGCCTCTCCCGCGGCTGGCGGATTCATTAATGACGTGGCACGACAGGTTT 1140
Db 3597 CGCAACCGCCTCTCCCGCGGCTGGCGGATTCATTAATGACGTGGCACGACAGGTTT 3538
QY 1141 CCGGACTGGAAGCGGCGAGTCAAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTAG 1200
Db 3537 CCGGACTGGAAGCGGCGAGTCAAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTAG 3478
QY 1201 GCAC 1204
Db 3477 GCAC 3474
```

RESULT 14

```
US-09-736-457-785/c
; Sequence 785, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
```

```
; CURRENT APPLICATION NUMBER: US/09/736.457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-785

Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGCGTATGGCATGATAGCCCGCGGAAGAGA 60
Db 4677 CGGACACCATCGAATGGCGCAAAACCTTTCCGCGTATGGCATGATAGCCCGCGGAAGAGA 4618
QY 61 GTCAATTTCAGGGTGGTGAATGTAACCAAGTAAAGTTATACGATGTCGACAGTATGCG 120
Db 4617 GTCAATTTCAGGGTGGTGAATGTAACCAAGTAAAGTTATACGATGTCGACAGTATGCG 4558
QY 121 GTGTCCTTTATCAGACCGGTTTCCGCGTGTGTAACCAAGTAAAGTTATACGATGTCG 180
Db 4557 GTGTCCTTTATCAGACCGGTTTCCGCGTGTGTAACCAAGTAAAGTTATACGATGTCG 4498
QY 181 CGCGGAAAAAGTGGAAAGCGCGCATGGCGAGCTGAATTAACATTTCCCAACCGCGTGGCAC 240
Db 4497 CGCGGAAAAAGTGGAAAGCGCGCATGGCGAGCTGAATTAACATTTCCCAACCGCGTGGCAC 4438
QY 241 AACAACTGGCGGGCAAAACAGTCGTTGTCGATTTGGCGTTGGCACCTTCAGTCTGGGCCCTGC 300
Db 4437 AACAACTGGCGGGCAAAACAGTCGTTGTCGATTTGGCGTTGGCACCTTCAGTCTGGGCCCTGC 4378
QY 301 ACGCGCGTCCGCAAAATTCGCGCGGATTAATCTCGCGCGCATCAACTGGGTGCCAGCG 360
Db 4377 ACGCGCGTCCGCAAAATTCGCGCGGATTAATCTCGCGCGCATCAACTGGGTGCCAGCG 4318
QY 361 TGGTGGTTCGATGGTAGAAGCAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCAACAATC 420
Db 4317 TGGTGGTTCGATGGTAGAAGCAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCAACAATC 4258
QY 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCAATTAACATTCGCTGGATGACAGGATGCCA 480
Db 4257 TTCTCGCGCAACGCGTCAGTGGGCTGATCAATTAACATTCGCTGGATGACAGGATGCCA 4198
QY 481 TTGCTGTGGAAGCTCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCCTGACCGAGA 540
Db 4197 TTGCTGTGGAAGCTCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCCTGACCGAGA 4138
QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGGCGGCCCATTAAGTTCTGTCTCGG 600
Db 4137 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGGCGGCCCATTAAGTTCTGTCTCGG 660
QY 601 TGGTCGCATTTGGTCAACAGCAAAATCGCGCTGTAGCGGCGCCATTAAGTTCTGTCTCGG 4018
Db 4077 TGGTCGCATTTGGTCAACAGCAAAATCGCGCTGTAGCGGCGCCATTAAGTTCTGTCTCGG 4018
QY 661 CGCGTCTGCGTCTGCGTGGCTGGCAATAATATCTCACTGCGCAATCAAAATTCAGCCGATAG 720
Db 4017 CGCGTCTGCGTCTGCGTGGCTGGCAATAATATCTCACTGCGCAATCAAAATTCAGCCGATAG 3958
QY 721 CGGAACGGGAAGGCACTGGAGTGGCATGTCGGTTCCTCAAAACCATGCAAAATGCTGA 780
Db 3957 CGGAACGGGAAGGCACTGGAGTGGCATGTCGGTTCCTCAAAACCATGCAAAATGCTGA 3898
QY 781 ATGAGGGCATGTTCCCACTCGGATGCTGGTTGCCAAGCATGAGATGGCGCTGGCGCAA 840
Db 3897 ATGAGGGCATGTTCCCACTCGGATGCTGGTTGCCAAGCATGAGATGGCGCTGGCGCAA 3838
QY 841 TCGCGGCCATTACCGAGTCCGGGCTCGCGTGGTGCGGATATCTCGGTAGTGGGATAG 900
Db 3837 TCGCGGCCATTACCGAGTCCGGGCTCGCGTGGTGCGGATATCTCGGTAGTGGGATAG 3778
```


QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACATCAAAACAGGATTTTC 960
Db 3777 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACATCAAAACAGGATTTTC 3718
QY 961 GCTGCTGGGGGAAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGGGTGA 1020
Db 3717 GCTGCTGGGGGAAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGGGTGA 3658
QY 1021 AGGGCAATCAGCTGTTGGCCGCTCTACTGTTGAAAGAAACACCCCTGGCGGCCAATA 1080
Db 3657 AGGGCAATCAGCTGTTGGCCGCTCTACTGTTGAAAGAAACACCCCTGGCGGCCAATA 3598
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGATTCAATTAATGACGTGGCACACAGGTTT 1140
Db 3597 CGCAAAACCGCTCTCCCGCGGTTGGCGATTCAATTAATGACGTGGCACACAGGTTT 3538
QY 1141 CCCGACTGGAAGCGGGAGTGAAGCGCAACGCAATTAATGAGTAGTCACTCAATTAG 1200
Db 3537 CCCGACTGGAAGCGGGAGTGAAGCGCAACGCAATTAATGAGTAGTCACTCAATTAG 3478
QY 1201 GCAC 1204
Db 3477 GCAC 3474

RESULT 15

US-09-614-124B-785/c
; Sequence 785, Application US/09614124B
; Patent No. 6630574

GENERAL INFORMATION:

; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darriek
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND

; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER

; FILE REFERENCE: 210121.478C9

; CURRENT APPLICATION NUMBER: US/09/614,124B

; CURRENT FILING DATE: 2001-07-11

; NUMBER OF SEQ ID NOS: 1668

; SOFTWARE: FactSeq for Windows Version 3.0

; SEQ ID NO 785

; LENGTH: 5502

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-614-124B-785

Query Match 36.3%; Score 1199.2; DB 4; Length 5502;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTCGGGTATGGCATGATACGGCCCGGAAGAGA 60
Db 4677 CGGACCATCGAATGGCGCAAAACCTTTCGGGTATGGCATGATACGGCCCGGAAGAGA 4618
QY 61 GTCAATTGAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGACAGATGTCGCG 120
Db 4617 GTCAATTGAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGACAGATGTCGCG 4558
QY 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACACAGGCGACCGATTTCTCGCAAAA 180
Db 4557 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACACAGGCGACCGATTTCTCGCAAAA 4498
QY 181 CGCGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTACATTCCCAACCGCGTGGCAC 240
Db 4497 CGCGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTACATTCCCAACCGCGTGGCAC 4438
QY 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGCGCACCTCCAGCTCGGCCCTGC 300

Search completed: August 3, 2005, 02:32:22
Job time : 356.4 secs

Db 4437 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 4378
QY 301 ACGGCGGTTCGCAAAATTTGTTCGCGCGGATTAATCTCGCGCGATCAACTGGGTCCGACGC 360
Db 4377 ACGGCGGTTCGCAAAATTTGTTCGCGCGGATTAATCTCGCGCGATCAACTGGGTCCGACGC 4318
QY 361 TGGTGGTGTGATGGTGAAGAGCGGCTCGAAGCCTGTGTAAGCGCGGTGCACAATC 420
Db 4317 TGGTGGTGTGATGGTGAAGAGCGGCTCGAAGCCTGTGTAAGCGCGGTGCACAATC 4258
QY 421 TTCTCGCGCAACCGTCAAGTGGGCTGATCAATTAATCTCGCGGATTAATCTCGCGGATGACCAATGCCA 480
Db 4257 TTCTCGCGCAACCGTCAAGTGGGCTGATCAATTAATCTCGCGGATTAATCTCGCGGATGACCAATGCCA 4198
QY 481 TTGCTGGAGAGTCCCTGCACATAATTTTCGCGGCTTAATTTCTTGATGTCTCTGACCAGA 540
Db 4197 TTGCTGGAGAGTCCCTGCACATAATTTTCGCGGCTTAATTTCTTGATGTCTCTGACCAGA 4138
QY 541 CACCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGATC 600
Db 4137 CACCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGATC 4078
QY 601 TGGTCGCAATTTGGGTCAACAGCAATCGCGCTTTAGCGGGCCCAATTAAGTTCTGTCGCG 660
Db 4077 TGGTCGCAATTTGGGTCAACAGCAATCGCGCTTTAGCGGGCCCAATTAAGTTCTGTCGCG 4018
QY 661 CGGCTGCGCTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAAAATCGTGA 720
Db 4017 CGGCTGCGCTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCGTGA 3958
QY 721 CGGAACGGGAAGCGGACTGGAGTGCCATGTCGCGTTTCAACAAACATGCAAAATGCTGA 780
Db 3957 CGGAACGGGAAGCGGACTGGAGTGCCATGTCGCGTTTCAACAAACATGCAAAATGCTGA 3898
QY 781 ATGAGGCGATCGTTCCCACTCGCATGCTGTTCCCAACGATCAGATGCGCTGGCGGCA 840
Db 3897 ATGAGGCGATCGTTCCCACTCGCATGCTGTTCCCAACGATCAGATGCGCTGGCGGCA 3838
QY 841 TGGCGCCATTACCGAGTCCGGCTCGGCTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 3837 TGGCGCCATTACCGAGTCCGGCTCGGCTGGTGGGATATCTCGGTAGTGGGATACG 3778
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACATCAAAACAGGATTTTC 960
Db 3777 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACATCAAAACAGGATTTTC 3718
QY 961 GCCTGCTGGGGCAAAACAGGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGGGTGA 1020
Db 3717 GCCTGCTGGGGCAAAACAGGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGGGTGA 3658
QY 1021 AGGGCAATCAGCTGTTGGCCGCTCTCACTGTTGAAGAAACCAACCCCTGGCGGCCAATA 1080
Db 3657 AGGGCAATCAGCTGTTGGCCGCTCTCACTGTTGAAGAAACCAACCCCTGGCGGCCAATA 3598
QY 1081 CGCAAAACCGCTCTCCCGCGGCTTGGCGGATTCATTAAATGAGTGGCACACAGGTTT 1140
Db 3597 CGCAAAACCGCTCTCCCGCGGCTTGGCGGATTCATTAAATGAGTGGCACACAGGTTT 3538
QY 1141 CCGGACTGGAAGCGGGAGTGAAGCGCAACGCAATTAATGAGTGGTCACTCAATTAG 1200
Db 3537 CCGGACTGGAAGCGGGAGTGAAGCGCAACGCAATTAATGAGTGGTCACTCAATTAG 3478
QY 1201 GCAC 1204
Db 3477 GCAC 3474

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 14:24:04 ; Search time 1094.6 Seconds
(without alignments)
17846.836 Million cell updates/sec

Title: US-09-765-555B-15
Perfect score: 3300
Sequence: 1 ccgacacatcgatggtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : N_Geneseq_16Dec04:*
- 1: Geneseqn1980s:*
 - 2: Geneseqn1990s:*
 - 3: Geneseqn2000s:*
 - 4: Geneseqn2001as:*
 - 5: Geneseqn2001bs:*
 - 6: Geneseqn2002as:*
 - 7: Geneseqn2002bs:*
 - 8: Geneseqn2003as:*
 - 9: Geneseqn2003bs:*
 - 10: Geneseqn2003cs:*
 - 11: Geneseqn2003ds:*
 - 12: Geneseqn2004as:*
 - 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3300	100.0	3300	4 AAD11588	Aad11588 Partial s
2	3300	100.0	3300	4 AAD11589	Aad11589 Partial s
3	3226.4	97.8	3300	4 AAD11592	Aad11592 Partial s
4	3215.2	97.4	3300	4 AAD11591	Aad11591 Partial s
5	3207.2	97.2	3300	4 AAD11590	Aad11590 Partial s
6	2705.8	82.0	8101	4 AAH25681	Aah25681 Nucleotid
7	2705.8	82.0	8101	6 ABL53238	Ab153238 Nucleotid
8	2690.4	81.5	7259	12 ADO23608	Ado23608 DNA encod
9	2690.4	81.5	7322	12 ADO23613	Ado23613 DNA encod
10	2690.4	81.5	7352	12 ADO23603	Ado23603 DNA encod
11	2690.4	81.5	7370	12 ADO23598	Ado23598 DNA encod
12	2690.4	81.5	7370	12 ADO23639	Ado23639 DNA encod
13	2690.4	81.5	7370	12 ADO23649	Ado23649 DNA encod
14	2690.4	81.5	7370	12 ADO23644	Ado23644 DNA encod
15	2690.4	81.5	7403	12 ADO23594	Ado23594 DNA encod
16	2690.4	81.5	7442	12 ADO23590	Ado23590 DNA encod
17	2690.4	81.5	7478	12 ADO23584	Ado23584 DNA encod
18	2688.8	81.5	7112	12 ADO23618	Ado23618 DNA encod
19	2636.8	79.9	6806	6 ABL49925	Ab149925 Maltose b
20	2605.4	79.0	7553	12 ADO23588	Ado23588 DNA encod

21	1609.4	48.8	5558	12 ADL90419	Adl90419 Clostridi
22	1474	44.7	4700	9 ACF06053	Adf06053 Vector pJ
23	1472.4	44.6	4700	12 ADL18670	Adl18670 Vector pJ
24	1462.4	44.3	4920	10 ADH73599	Adh73599 Plasmid p
25	1462.4	44.3	4935	10 ADH73598	Adh73598 Plasmid p
26	1462.4	44.3	4945	10 ADH73600	Adh73600 Plasmid p
27	1462.4	44.3	4951	10 ADH73601	Adh73601 Plasmid p
28	1418.2	43.0	5903	8 ABZ23939	Abz23939 Nucleotid
29	1417.6	43.0	1511	6 ABQ73200	Abq73200 E. coli t
30	1417.6	43.0	5926	2 AAV32977	Aav32977 Tn7 donor
31	1417.6	43.0	5926	6 AAD45059	Aad45059 Transposo
32	1417.6	43.0	5926	12 ADG46817	Adg46817 Donor pla
33	1362	41.3	5201	12 ADL72229	Adl72229 DNA seque
34	1362	41.3	5201	12 ADL72228	Adl72228 DNA seque
35	1238.4	37.5	5024	9 ACF06051	Adf06051 Plasmid p
36	1238.4	37.5	5024	12 ADL18668	Adl18668 Plasmid p
37	1237.2	37.5	1922	6 ABQ73089	Abq73089 MBP and z
38	1237.2	37.5	1922	10 ADF83622	Adf83622 MBP-human
39	1204	36.5	4557	2 AAT90491	Aat90491 Vector pI
40	1204	36.5	4969	12 ADO33522	Ado33522 Expressio
41	1204	36.5	5006	10 ADE24112	Ade24112 plasmid p
42	1204	36.5	5064	12 ADO33532	Ado33532 Vector pG
43	1204	36.5	5082	12 ADO33531	Ado33531 Vector pG
44	1204	36.5	6259	6 AAD29720	Aad29720 Plasmid p
45	1204	36.5	6823	3 AAC55459	Aac55459 Destinati

ALIGNMENTS

RESULT 1	
AAD11588	
ID	AAD11588 standard; DNA; 3300 BP.
XX	
AC	AAD11588;
XX	
DT	24-SEP-2001 (first entry)
XX	
DE	Partial sequence of pMal-m1 and ZFPm1 DNA.
XX	
KW	Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
KW	modulation; plant technology; agriculture; ds.
XX	
OS	Unidentified.
XX	
PH	Key
FT	CDS
FT	Location/Qualifiers
FT	2719..3270
FT	/tag=a
FT	/product="ZFPm1 protein"
FT	/transl_except= (pos:2785..2787, aa:Asp)
FT	/transl_except= (pos:2788..2790, aa:Pro)
FT	/transl_except= (pos:2791..2793, aa:Gly)
FT	/transl_except= (pos:2872..2874, aa:Arg)
FT	/transl_except= (pos:2875..2877, aa:Ala)
FT	/transl_except= (pos:2878..2880, aa:His)
FT	/transl_except= (pos:2884..2886, aa:Glu)
FT	/transl_except= (pos:2953..2955, aa:Gln)
FT	/transl_except= (pos:2959..2961, aa:Ser)
FT	/transl_except= (pos:3121..3123, aa:Arg)
FT	/transl_except= (pos:3124..3126, aa:Ser)
FT	/transl_except= (pos:3127..3129, aa:Asp)
FT	/transl_except= (pos:3130..3132, aa:Asn)
FT	/transl_except= (pos:3136..3138, aa:Val)
FT	/transl_except= (pos:3139..3141, aa:Arg)
FT	/transl_except= (pos:3205..3207, aa:Gln)
FT	/transl_except= (pos:3208..3210, aa:Ala)
FT	/transl_except= (pos:3211..3213, aa:Gly)
FT	/transl_except= (pos:3214..3216, aa:His)
FT	/transl_except= (pos:3220..3222, aa:Ala)
FT	/transl_except= (pos:3223..3225, aa:Ser)
FT	/note="CDS does not include start and stop codon"
FT	/partial
FT	primer_bind 2740..2790

```

FT primer_bind /tag= b
FT /bound_moiety= "F1-f2 primer"
FT 2770..2850
FT /tag= c
FT primer_bind /bound_moiety= "F1-f1 primer"
FT 2824..2889
FT /tag= d
FT primer_bind /bound_moiety= "F2-b primer"
FT 2867..2940
FT /tag= e
FT primer_bind /bound_moiety= "F2-f primer"
FT 2916..2973
FT /tag= f
FT primer_bind /bound_moiety= "F3-b1 primer"
FT 2953..3021
FT /tag= g
FT primer_bind /bound_moiety= "F3-b2 primer"
FT 2992..3042
FT /tag= h
FT primer_bind /bound_moiety= "F4-f2 primer"
FT 3022..3102
FT /tag= i
FT primer_bind /bound_moiety= "F4-f1 primer"
FT 3076..3141
FT /tag= j
FT primer_bind /bound_moiety= "F5-b primer"
FT 3119..3192
FT /tag= k
FT primer_bind /bound_moiety= "F5-f primer"
FT 3168..3225
FT /tag= l
FT primer_bind /bound_moiety= "F6-b1 primer"
FT 3205..3273
FT /tag= m
FT primer_bind /bound_moiety= "F6-b2 primer"

```

WO200152620-A2.

26-JUL-2001.

19-JAN-2001; 2001WO-US001817.

21-JAN-2000; 2000US-0177468P.

21-JUL-2000; 2000US-00620897.

(SCRI) SCRIPPS RES INST.

(SYGN) SYNGENTA AGRIC DISCOVERY INC.

Barbas CF, Stege JT, Guan X, Dalmia B;

WPI; 2001-465325/50.

P-PSDB; AAE06000.

New zinc finger proteins, useful for modulating or regulating gene expression and metabolic pathways in plants, e.g. for treating in the plant cells a disorder that is associated with abnormal expression of the target gene.

Example 4; Page 138-139; 156pp; English.

The patent discloses methods and compositions to modulate the expression of a target gene in plant cells. The method involves providing plant cells with a zinc finger protein (ZFP) which is capable of specifically binding to a target nucleotide sequence or its complementary strand within a target gene and allowing the ZFP binding to the target nucleotide sequence, where the expression of the target gene in the plant cells is modulated. The ZFP and fusions of the ZFP proteins are useful for modulating or regulating gene expression and metabolic pathways in plants. The ZFP, fusion proteins and methods are useful in plant and agricultural technology. The method is useful particularly for treating a disorder in the plant cells, where the disorder is associated with abnormal expression of the target gene. The present DNA sequence is the partial sequence of pMal-m1 and ZFPm1 DNA

```

XX SQ Sequence 3300 BP; 847 A; 874 C; 876 G; 703 T; 0 U; 0 Other;
Query Match 100.0%; Score 3300; DB 4; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCGACACATCGAATGTTGCAAAACCTTTTCGGGTATGGCATGATACGCCCGGAAGAGA 60
DB 1 CCGACACATCGAATGTTGCAAAACCTTTTCGGGTATGGCATGATACGCCCGGAAGAGA 60
QY 61 GTCAATTCAGGGTGTGAATGTGAACACAGTAACTTATACGATGTGCGAGAGTATCCG 120
DB 61 GTCAATTCAGGGTGTGAATGTGAACACAGTAACTTATACGATGTGCGAGAGTATCCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACACAGGCGAGCCAGCTTCTGCGAAAA 180
DB 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACACAGGCGAGCCAGCTTCTGCGAAAA 180
QY 181 CCGGGGAAAAGTGAAGCGCGGATGCGGAGCTGAATTAATTCACACCGGTGGCAC 240
DB 181 CCGGGGAAAAGTGAAGCGCGGATGCGGAGCTGAATTAATTCACACCGGTGGCAC 240
QY 241 AACAACTGGCGGGCAAAACAGTCTGCTGATTGCGCGTTGCCACTCCAGTCTGSCCTGC 300
DB 241 AACAACTGGCGGGCAAAACAGTCTGCTGATTGCGCGTTGCCACTCCAGTCTGSCCTGC 300
QY 301 ACGCGCGTGCAGAAATTTGTCGCGCGATTAATCTCGCGCGGATCACTGGGTGCCAGCG 360
DB 301 ACGCGCGTGCAGAAATTTGTCGCGCGATTAATCTCGCGCGGATCACTGGGTGCCAGCG 360
QY 361 TGGTGTGTCGATGTTAGAACGAGCGGCTGCAAGCTGTAAAGCGCGGTGCAATC 420
DB 361 TGGTGTGTCGATGTTAGAACGAGCGGCTGCAAGCTGTAAAGCGCGGTGCAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACACAGATGCCA 480
DB 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACACAGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACATAATTTCCGCGCTTAATTTCTTGATGTCTCTGACCAG 540
DB 481 TTGCTGTGGAAGCTGCTGCACATAATTTCCGCGCTTAATTTCTTGATGTCTCTGACCAG 540
QY 541 CACCATCAACAGTATTATTTCTCCCATGAAGCGGTACGCGACTGGGGGTGGAGCATC 600
DB 541 CACCATCAACAGTATTATTTCTCCCATGAAGCGGTACGCGACTGGGGGTGGAGCATC 600
QY 601 TGGTTCGATTTGGTCCACAGCAAAATCGCGTGTAGCGGGGCCATTAAGTTCTGTCTCGG 660
DB 601 TGGTTCGATTTGGTCCACAGCAAAATCGCGTGTAGCGGGGCCATTAAGTTCTGTCTCGG 660
QY 661 CCGGTCTGGCTTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CCGGTCTGGCTTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CCGAAGCGGAAGCGGACTGGAGTGCCATGTCGGTTTTTCAACAAACCATGCAATCTGA 780
DB 721 CCGAAGCGGAAGCGGACTGGAGTGCCATGTCGGTTTTTCAACAAACCATGCAATCTGA 780
QY 781 ATGAGGCGCATCGTTCCACTCGCATGCTGGTTGCCAACGATCAGATGGGCGTGGGCGCAA 840
DB 781 ATGAGGCGCATCGTTCCACTCGCATGCTGGTTGCCAACGATCAGATGGGCGTGGGCGCAA 840
QY 841 TGGCGCGCATTAACGAGTCCGGGTGCGGTGGTGGGATATCTCGGTAGTGGGATAG 900
DB 841 TGGCGCGCATTAACGAGTCCGGGTGCGGTGGTGGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACACCATCAAAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACACCATCAAAACAGGATTTTC 960
QY 961 GCTTGTGCGGGCAAAACAGCGTGGACCGCTTCTGCTCAACTCTCTCAGGCGCCAGCGGTGA 1020
DB 961 GCTTGTGCGGGCAAAACAGCGTGGACCGCTTCTGCTCAACTCTCTCAGGCGCCAGCGGTGA 1020

```

Db 961 GCCTGCTGGGCAACACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGCCAGGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTGCGCGTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGCGCGTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCCTCTCCCGCGGTGGCCGATTCAATGAGCTGGCGACGACAGGTTT 1140
Db 1081 CGCAACCGCCTCTCCCGCGGTGGCCGATTCAATGAGCTGGCGACGACAGGTTT 1140
Qy 1141 CCGGACTGGAAGCGGAGTGAAGCGCAACCAATTAATGTAGTCTCACTCATTAG 1200
Db 1141 CCGGACTGGAAGCGGAGTGAAGCGCAACCAATTAATGTAGTCTCACTCATTAG 1200
Qy 1201 GCACAATCTCATGTTTGAAGCTATCATGCACTGCAAGTGCACCAATGTTTGGCG 1260
Db 1201 GCACAATCTCATGTTTGAAGCTATCATGCACTGCAAGTGCACCAATGTTTGGCG 1260
Qy 1261 TCAGGAGCCATCGGAAGCTGGTATGGCTGTGAGGTCGTAATCACTGATTAATTCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGGTATGGCTGTGAGGTCGTAATCACTGATTAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTCTGGATAATGTTTTTGGCGGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTGGATAATGTTTTTGGCGGACATCAATACGGTT 1380
Qy 1381 CTGGCAATATTTCTGAATGAGCTGTGACAATTAATCAATCGGCTCGTATAATGTGGA 1440
Db 1381 CTGGCAATATTTCTGAATGAGCTGTGACAATTAATCAATCGGCTCGTATAATGTGGA 1440
Qy 1441 ATTGTGAGCGGATACAAATTTACACAGGAACAGCCAGTCCGTTAGTGTTTTTCAG 1500
Db 1441 ATTGTGAGCGGATACAAATTTACACAGGAACAGCCAGTCCGTTAGTGTTTTTCAG 1500
Qy 1501 GCACCTTACCAACAGGACCATAGATTATGAAGAACTGAAGAGGTAAACTGGTAACTCG 1560
Db 1501 GCACCTTACCAACAGGACCATAGATTATGAAGAACTGAAGAGGTAAACTGGTAACTCG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAG 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAG 1620
Qy 1621 ACCGNAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGNAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTATCTTGGGACACGACCGCTTTGGTGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTATCTTGGGACACGACCGCTTTGGTGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGCTGAATTCACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAATTCACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGATGCCGTAGCTTACACGGCAAGCTGATTCCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCCGTAGCTTACACGGCAAGCTGATTCCTTACCCGATCGCTGTT 1860
Qy 1861 GAAGCGTATCGTGAATTTATACAAAGATCTGTCGCCAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTATCGTGAATTTATACAAAGATCTGTCGCCAACCCGCCAAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTATCTTCACTGGCGCGCTGATGCTGCTGAACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTATCTTCACTGGCGCGCTGATGCTGCTGAACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAACGGCAAGTACCACTTAAGACGTTGGCGTGGATAACCGCTGGCGGAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACCACTTAAGACGTTGGCGTGGATAACCGCTGGCGGAAGCG 2100

Qy 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAAACAAACACATGAATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAAACAAACACATGAATGACAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAAAGCTGCCCTTTAATAAAGCGGAAACAGCGATGACCATCAAGCGCCCGTGG 2220
Db 2161 TCCATCGCAGAAAGCTGCCCTTTAATAAAGCGGAAACAGCGATGACCATCAAGCGCCCGTGG 2220
Qy 2221 GCATGGTCCAAACATCGACACAGCAAAAGTGAATTAATGGTTAAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAACATCGACACAGCAAAAGTGAATTAATGGTTAAACGGTACTGCGGACCTTC 2280
Qy 2281 AAGGTCAAACATCCAAACCGTTTGGTGGCTGCTGAGCGCAGGTATTAAACGCCGCCAGT 2340
Db 2281 AAGGTCAAACATCCAAACCGTTTGGTGGCTGCTGAGCGCAGGTATTAAACGCCGCCAGT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTCACTGATGAAGGCTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTCACTGATGAAGGCTCTG 2400
Qy 2401 GAAGCGGTTAATAAAGACAAACCGCTGGTGCCTGAGCGCTGAAGTCTTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCGCTGGTGCCTGAGCGCTGAAGTCTTTACGAGGAAG 2460
Qy 2461 TTGGCAAAAGATCCACGTATTGCGGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCAAAAGATCCACGTATTGCGGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTCTGGTATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTCTGGTATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Qy 2581 GCACGCGTCTGACAGTCTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCACGCGTCTGACAGTCTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Qy 2641 AACAAACAACAATAAACAATAAACAACCTCGGGATCGAGGGAAGATTTCAGAAATTC 2700
Db 2641 AACAAACAACAATAAACAATAAACAACCTCGGGATCGAGGGAAGATTTCAGAAATTC 2700
Qy 2701 GGATCCTCTCTCTGTCGCGCCAGCGGCTTCGAGCCCGGGGAGAGCCCTATGCTGT 2760
Db 2701 GGATCCTCTCTCTGTCGCGCCAGCGGCTTCGAGCCCGGGGAGAGCCCTATGCTGT 2760
Qy 2761 CCGGAATGTGTAAGTCTCTCTCAGAGCTCTCACTGCTGCGCACCCAGCGTACCCAC 2820
Db 2761 CCGGAATGTGTAAGTCTCTCTCAGAGCTCTCACTGCTGCGCACCCAGCGTACCCAC 2820
Qy 2821 ACGGTGAAAAACCGTATAAATGCCAGATGCGGCAAACTTTTATGCGCAGTCCAGCAAC 2880
Db 2821 ACGGTGAAAAACCGTATAAATGCCAGATGCGGCAAACTTTTATGCGCAGTCCAGCAAC 2880
Qy 2881 CTGGTCCGCATCAACGCACTCTACTGGGAGAGCCATAAATGTCCAGAAATGTTGGC 2940
Db 2881 CTGGTCCGCATCAACGCACTCTACTGGGAGAGCCATAAATGTCCAGAAATGTTGGC 2940
Qy 2941 AAGTCTTTCTCTCGGTGACAAATCTCGTCCGGCACCAACGTACTCACACCGGGGAGAG 3000
Db 2941 AAGTCTTTCTCTCGGTGACAAATCTCGTCCGGCACCAACGTACTCACACCGGGGAGAG 3000
Qy 3001 CCCTATGCTTGTCCGGAATGTTAAGTCTTTCAGCCGAGCGATAAACCCTGGTGGGCCAC 3060
Db 3001 CCCTATGCTTGTCCGGAATGTTAAGTCTTTCAGCCGAGCGATAAACCCTGGTGGGCCAC 3060
Qy 3061 CAGCGTACCCACACGGGTGAAGAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACGGGTGAAGAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTAGC 3120
Qy 3121 CAGGCGGCGCACTGGCGAGCCATCAACGCACTCTACTGCGGAGAGCCATACAAATGT 3180
Db 3121 CAGGCGGCGCACTGGCGAGCCATCAACGCACTCTACTGCGGAGAGCCATACAAATGT 3180

QY 3181 CCAGAAATGGCAAGTCTTTCTCTCGGTCTGCAATCTCGTCCGGCACCACCAAGTACTAC 3240
 Db |||||
 3181 CCAGAAATGGCAAGTCTTTCTCTCGGTCTGCAATCTCGTCCGGCACCACCAAGTACTAC 3240
 QY 3241 ACCGGTAAAAAATAGTGGCCAGGCGCGCGAGTACCGGTACGACGTTCCGGACTACGCT 3300
 Db |||||
 3241 ACCGGTAAAAAATAGTGGCCAGGCGCGCGAGTACCGGTACGACGTTCCGGACTACGCT 3300

RESULT 2

AAD11589

ID AAD11589 standard; DNA; 3300 BP.

XX AC AAD11589;

XX DT 24-SEP-2001 (first entry)

XX DE Partial sequence of pMal-m2 and ZFPm2 DNA.

XX KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;

XX KW modulation; plant technology; agriculture; ds.

XX OS Unidentified.

XX FH Key

FT CDS

FT Location/Qualifiers

FT 2719..3270

FT /*tag= a

FT /product= "ZFPm2 protein"

FT /note= "CDS does not include start and stop codon"

FT /partial

FT primer_bind

FT 2740..2790

FT /*tag= b

FT /bound_moiety= "F1-f2 primer"

FT 2770..2850

FT /*tag= c

FT /bound_moiety= "F1-f1 primer"

FT 2824..2889

FT /*tag= d

FT /bound_moiety= "F2-b primer"

FT 2867..2940

FT /*tag= e

FT /bound_moiety= "F2-f primer"

FT 2916..2973

FT /*tag= f

FT /bound_moiety= "F3-b1 primer"

FT 2953..3021

FT /*tag= g

FT /bound_moiety= "F3-b2 primer"

FT 2992..3042

FT /*tag= h

FT /bound_moiety= "F4-f2 primer"

FT 3022..3102

FT /*tag= i

FT /bound_moiety= "F4-f1 primer"

FT 3076..3141

FT /*tag= j

FT /bound_moiety= "F5-b primer"

FT 3119..3192

FT /*tag= k

FT /bound_moiety= "F5-f primer"

FT 3168..3225

FT /*tag= l

FT /bound_moiety= "F6-b1 primer"

FT 3205..3273

FT /*tag= m

FT /bound_moiety= "F6-b2 primer"

XX WO200152620-A2.

XX PN 26-JUL-2001.

XX PD 19-JAN-2001; 2001WO-US001817.

XX PF

XX

XX

XX

XX

PR

PR 21-JAN-2000; 2000US-0177468P.

XX 21-JUL-2000; 2000US-00620897.

PA (SCRI) SCRIPPS RES INST.

PA (SYGN) SYNGENTA AGRIC DISCOVERY INC.

XX

PI Barbas CF, Stege JT, Guan X, Dalmia B;

XX WPI; 2001-465325/50.

DR P-PSDB; RAE06002.

XX

XX New zinc finger proteins, useful for modulating or regulating gene

PT expression and metabolic pathways in plants, e.g. for treating in the

PT plant cells a disorder that is associated with abnormal expression of the

PT target gene.

XX

XX Example 4; Page 140-142; 156pp; English.

XX

CC The patent discloses methods and compositions to modulate the expression

CC of a target gene in plant cells. The method involves providing plant

CC cells with a zinc finger protein (ZFP) which is capable of specifically

CC binding to a target nucleotide sequence or its complementary strand

CC within a target gene and allowing the ZFP binding to the target

CC nucleotide sequence. Where the expression of the target gene in the plant

CC cells is modulated. The ZFP and fusions of the ZFP proteins are useful

CC for modulating or regulating gene expression and metabolic pathways in

CC plants. The ZFP, fusion proteins and methods are useful in plant and

CC agricultural technology. The method is useful particularly for treating a

CC disorder in the plant cells, where the disorder is associated with

CC abnormal expression of the target gene. The present DNA sequence is the

CC partial sequence of pMal-m2 and ZFPm2 DNA

XX

SQ Sequence 3300 BP; 847 A; 874 C; 876 G; 703 T; 0 U; 0 Other;

Query Match

Best Local Similarity 100.0%; Score 3300; DB 4; Length 3300;

Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGAAGAGA 60

Db |||||

1 CCACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGAAGAGA 60

QY 61 GTCATTCAGGGTGGTGAATGTGAACACAGTACGTTATACGATGTCGACAGATATGCCG 120

Db |||||

61 GTCATTCAGGGTGGTGAATGTGAACACAGTACGTTATACGATGTCGACAGATATGCCG 120

QY 121 GTGTCTTTATCAGACCGTTTCCCGGTGTGAACACAGCCACCGTTCCTTCGCAAAA 180

Db |||||

121 GTGTCTTTATCAGACCGTTTCCCGGTGTGAACACAGCCACCGTTCCTTCGCAAAA 180

QY 181 CGCGGCAAAAAGTGAAGCGCGGATCGCGAGCTGAATTACATTTCCCAACCGGTGGCAC 240

Db |||||

181 CGCGGCAAAAAGTGAAGCGCGGATCGCGAGCTGAATTACATTTCCCAACCGGTGGCAC 240

QY 241 AACAACTGGCGGCAAAACAGTCTGTTGATTTGGCGTTTGCACCTCCAGTCTGGCCCTGC 300

Db |||||

241 AACAACTGGCGGCAAAACAGTCTGTTGATTTGGCGTTTGCACCTCCAGTCTGGCCCTGC 300

QY 301 ACGCGCGTCCGCAAAATTGTCCGGCGGATTAATCTCGCGCCGATCAACTGGGTGCGAGCG 360

Db |||||

301 ACGCGCGTCCGCAAAATTGTCCGGCGGATTAATCTCGCGCCGATCAACTGGGTGCGAGCG 360

QY 361 TGGTGTGTTCGATGGTAGAACGAAGCGCGTTCGAAGCCTGTAAAGCGCGGTGCACAATC 420

Db |||||

361 TGGTGTGTTCGATGGTAGAACGAAGCGCGTTCGAAGCCTGTAAAGCGCGGTGCACAATC 420

QY 421 TTCTCGCGCAACCGCTCAGTGGGTGATCATTAATCATCCGCTGGATGACAGGATGCCA 480

Db |||||

421 TTCTCGCGCAACCGCTCAGTGGGTGATCATTAATCATCCGCTGGATGACAGGATGCCA 480

QY 481 TTGCTGTGAAGCTCCTCAGTAAATGTTCCGGCGTTATTCTTCTGATGTCCTGACCAGA 540

Db |||||

481 TTGCTGTGAAGCTCCTCAGTAAATGTTCCGGCGTTATTCTTCTGATGTCCTGACCAGA 540

QY	541	CACCCATCAACAGTATTATTTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTATTTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600
QY	601	TGGTCGCATTTGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG	660
DB	601	TGGTCGCATTTGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG	660
QY	661	CGCGTCTGCGTCTGGCTGGCGTAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	661	CGCGTCTGCGTCTGGCTGGCGTAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAACGGGAAGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA	780
DB	721	CGGAACGGGAAGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA	780
QY	781	ATGAGGGCATCGTTCCTCCACTGCGATGCTGTGGTTCGCAACGATCGATGGCGTGGCGCAA	840
DB	781	ATGAGGGCATCGTTCCTCCACTGCGATGCTGTGGTTCGCAACGATCGATGGCGTGGCGCAA	840
QY	841	TGCGGCCATTAACCGACTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG	900
DB	841	TGCGGCCATTAACCGACTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAAAGACAGCTCATGTTATATCCCGCGTTTAAACCAATCAAAACAGGATTTTC	960
DB	901	ACGATACCGAAAGACAGCTCATGTTATATCCCGCGTTTAAACCAATCAAAACAGGATTTTC	960
QY	961	GCCTGCTGGGCAACACGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGCGGTGA	1020
DB	961	GCCTGCTGGGCAACACGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGCGGTGA	1020
QY	1021	AGGGCAATCAGCTGTTGCGCGTCTCACTGGTGAAAGAAAACCAACCTGGCGGCCAATA	1080
DB	1021	AGGGCAATCAGCTGTTGCGCGTCTCACTGGTGAAAGAAAACCAACCTGGCGGCCAATA	1080
QY	1081	CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGTGGCGACGACAGGTTT	1140
DB	1081	CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGTGGCGACGACAGGTTT	1140
QY	1141	CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG	1200
DB	1141	CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG	1200
QY	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAGGTGCACCAATGCTCTGCGG	1260
DB	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAGGTGCACCAATGCTCTGCGG	1260
QY	1261	TCAGGCAGCCATCGGAAGCTGGTATGGCTGTGCGAGTTCGTAATCACTGCGATTAATTCG	1320
DB	1261	TCAGGCAGCCATCGGAAGCTGGTATGGCTGTGCGAGTTCGTAATCACTGCGATTAATTCG	1320
QY	1321	TGTCGCTCAAGCGCACTCCCGTCTGGATATGTTTTTGGCGCGACATCATACGGTT	1380
DB	1321	TGTCGCTCAAGCGCACTCCCGTCTGGATATGTTTTTGGCGCGACATCATACGGTT	1380
QY	1381	CTGGCAATATTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTGTGA	1440
DB	1381	CTGGCAATATTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTGTGA	1440
QY	1441	ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTTAGGTGTTTTCAAGA	1500
DB	1441	ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTTAGGTGTTTTCAAGA	1500
QY	1501	GCACCTTACCACACAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAAATCTGG	1560
DB	1501	GCACCTTACCACACAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAAATCTGG	1560
QY	1561	ATTTAAGCGGATAAAGGCTATACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT	1620
DB	1561	ATTTAAGCGGATAAAGGCTATACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT	1620

QY	1621	ACCGGAATTAAGTCAACGTTGAGCATCCGGATAAACTGGAGAGAAATTTCCACAGGTT	1680
DB	1621	ACCGGAATTAAGTCAACGTTGAGCATCCGGATAAACTGGAGAGAAATTTCCACAGGTT	1680
QY	1681	CGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
DB	1681	CGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCCTGTTGGTGAATCAACCCGGACAAAGCGTTCAGGACAAGCTGTAT	1800
DB	1741	GCTCAATCTGGCCTGTTGGTGAATCAACCCGGACAAAGCGTTCAGGACAAGCTGTAT	1800
QY	1801	CCGTTTACCTGGGATGCGGTACGTTAACAAGCAAGCTGATTGCTTACCCGATCGCTGT	1860
DB	1801	CCGTTTACCTGGGATGCGGTACGTTAACAAGCAAGCTGATTGCTTACCCGATCGCTGT	1860
QY	1861	GAAGCGTTATCGCTGATTATAACAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA	1920
DB	1861	GAAGCGTTATCGCTGATTATAACAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA	1920
QY	1921	GAGATCCCGGCGCTGGATAAAGAACTGAAGCGAAAGTAAAGCGCGCTGATTTCAAC	1980
DB	1921	GAGATCCCGGCGCTGGATAAAGAACTGAAGCGAAAGTAAAGCGCGCTGATTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTGCGCGTGTATGCTGCTGACGGGGGTTATGCGTTCAAG	2040
DB	1981	CTGCAAGAACCGTACTTCACTGCGCGTGTATGCTGCTGACGGGGGTTATGCGTTCAAG	2040
QY	2041	TATGAAGCGCAAGTACGACATTAAGAGCTGGCGGTGGATTAACGCTGGCGCGAAGCG	2100
DB	2041	TATGAAGCGCAAGTACGACATTAAGAGCTGGCGGTGGATTAACGCTGGCGCGAAGCG	2100
QY	2101	GSTCTGACCTTCTGCTGCTGACCTGATTAAAGCAACACATGAATGCAGACACCGATTAC	2160
DB	2101	GSTCTGACCTTCTGCTGCTGACCTGATTAAAGCAACACATGAATGCAGACACCGATTAC	2160
QY	2161	TCCATCGCAAGAGTGCCTTTTAAAGGCGGAAACAGCGATGACCATCAACGGCCCGTGG	2220
DB	2161	TCCATCGCAAGAGTGCCTTTTAAAGGCGGAAACAGCGATGACCATCAACGGCCCGTGG	2220
QY	2221	GCATGTCACACATCGACACCGCAAGTGAATTAATGGTGAACCGTACTGCGGACCTTC	2280
DB	2221	GCATGTCACACATCGACACCGCAAGTGAATTAATGGTGAACCGTACTGCGGACCTTC	2280
QY	2281	AAGGTCACACCATCAAAACCGTTCTGAGCGCGTGTGAGCGCAGGTATTAAACGCCACGT	2340
DB	2281	AAGGTCACACCATCAAAACCGTTCTGAGCGCGTGTGAGCGCAGGTATTAAACGCCACGT	2340
QY	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCGTG	2400
DB	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCGTG	2400
QY	2401	GAAGCGTTTAAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	2460
DB	2401	GAAGCGTTTAAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT	2460
QY	2461	TTGGCGAAAGATCCACGATTTGCGGCCACCATGGAACAAACCGCCAGAAAGTGAATCATG	2520
DB	2461	TTGGCGAAAGATCCACGATTTGCGGCCACCATGGAACAAACCGCCAGAAAGTGAATCATG	2520
QY	2521	CCGAACATCCCGAGATGTCGCGTTTTCTGGTATGTCGCGTGGTATGTCGGGTGATCAACGCC	2580
DB	2521	CCGAACATCCCGAGATGTCGCGTTTTCTGGTATGTCGCGTGGTATGTCGGGTGATCAACGCC	2580
QY	2581	GCCAGCGTCTGACGCTGATGAAGCGCTGGAAGACGCGCAGACTTAATTCGAGCTCG	2640
DB	2581	GCCAGCGTCTGACGCTGATGAAGCGCTGGAAGACGCGCAGACTTAATTCGAGCTCG	2640
QY	2641	AACAACAAACAAATAACAAACAAACCTCGGGATCGAGGGAAGGATTTTCAAAATTC	2700
DB	2641	AACAACAAACAAATAACAAACAAACCTCGGGATCGAGGGAAGGATTTTCAAAATTC	2700
QY	2701	GGATCCTCTTCTGCTGGCCCAAGCGCGCTCTCGAGCCCGGGGAGAACCCCTATGCTGTG	2760


```
Db 2701 GGATCCTCTCTCTGTGTGGCCAGCGGCGCTTCGAGCCGGGAGAACCTTATGCTTGT 2760
QY 2761 CCGGAATGTGTAAGTCTCTCTCAGAGCTCTACCTGGTGGCCACAGCGTACCCAC 2820
Db 2761 CCGGAATGTGTAAGTCTCTCTCAGAGCTCTACCTGGTGGCCACAGCGTACCCAC 2820
QY 2821 ACGGCTGAAAAACCGTATATAATGCCAGAGTGGCGCAAAATCTTTTAGCAGTCCAGCAAC 2880
Db 2821 ACGGCTGAAAAACCGTATATAATGCCAGAGTGGCGCAAAATCTTTTAGCAGTCCAGCAAC 2880
QY 2881 CTGGTGGCCCATCAACGCACTCATATCTGGCGAGAGCCATCAAAATGTCCAGAATGTGGC 2940
Db 2881 CTGGTGGCCCATCAACGCACTCATATCTGGCGAGAGCCATCAAAATGTCCAGAATGTGGC 2940
QY 2941 AAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTACTCACACCGGGGAGAG 3000
Db 2941 AAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTACTCACACCGGGGAGAG 3000
QY 3001 CCTATGCTTGTCCGGAATGTGGTAAGTCTTTCAGCCGCGAGGATAACCTGGTGGCCAC 3060
Db 3001 CCTATGCTTGTCCGGAATGTGGTAAGTCTTTCAGCCGCGAGGATAACCTGGTGGCCAC 3060
QY 3061 CAGCGTACCCACACGGGTGAAAAACCGTATATAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACGGGTGAAAAACCGTATATAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
QY 3121 CAGGCGGCGCACCTGGCGAGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
Db 3121 CAGGCGGCGCACCTGGCGAGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
QY 3181 CCAGAATGTGCAAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTACTCAC 3240
Db 3181 CCAGAATGTGCAAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTACTCAC 3240
QY 3241 ACCGGTAAAAAATAGTGGCCAGCGCGCGCCAGTACCGGTACGAGTTCGGGACTAGCCT 3300
Db 3241 ACCGGTAAAAAATAGTGGCCAGCGCGCGCCAGTACCGGTACGAGTTCGGGACTAGCCT 3300

RESULT 3
ID AAD11592 standard; DNA; 3300 BP.
XX AC AAD11592;
XX DT 24-SEP-2001 (first entry)
XX DE Partial sequence of pMal- $\alpha$ 3 and ZFP $\alpha$ 3 DNA.
XX KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
XX KW modulation; plant technology; agriculture;  $\alpha$ 3; APETALA3; ds.
XX OS Unidentified.
XX PH Key
XX FT Location/Qualifiers
XX CDS 2719..3270
XX /tag= a
XX /product= "ZFP $\alpha$ 3 protein"
XX /transl_except= (pos:3124..3126, aa:Ser)
XX /transl_except= (pos:3127..3129, aa:Ser)
XX /transl_except= (pos:3136..3138, aa:Val)
XX /note= "CDS does not include start and stop codon"
XX /partial
XX primer_bind 2740..2790
XX /tag= b
XX /bound_moiety= "F1-f2 primer"
XX primer_bind 2770..2850
XX /tag= c
XX /bound_moiety= "F1-f1 primer"
XX primer_bind 2824..2889
XX /tag= d
XX /bound_moiety= "F2-b primer"
```

```
FT primer_bind 2867..2940
FT /tag= e
FT /bound_moiety= "F2-f primer"
FT primer_bind 2916..2973
FT /tag= f
FT /bound_moiety= "F3-b1 primer"
FT primer_bind 2953..3021
FT /tag= g
FT /bound_moiety= "F3-b2 primer"
FT primer_bind 2992..3042
FT /tag= h
FT /bound_moiety= "F4-f2 primer"
FT primer_bind 3022..3102
FT /tag= i
FT /bound_moiety= "F4-f1 primer"
FT primer_bind 3076..3141
FT /tag= j
FT /bound_moiety= "F5-b primer"
FT primer_bind 3119..3192
FT /tag= k
FT /bound_moiety= "F5-f primer"
FT primer_bind 3168..3225
FT /tag= l
FT /bound_moiety= "F6-b1 primer"
FT primer_bind 3205..3273
FT /tag= m
FT /bound_moiety= "F6-b2 primer"
XX
XX W0200152620-A2.
XX
XX 26-JUL-2001.
XX
XX 19-JAN-2001; 2001WO-US001817.
XX
XX 21-JAN-2000; 2000US-0177468P.
XX
XX 21-JUL-2000; 2000US-00620897.
XX
XX (SCRI ) SCRIPPS RES INST.
XX (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
XX Barbas CF, Stege JT, Guan X, Dalmia B;
XX
XX WPI; 2001-465325/50.
XX P-PSDB; AAE06005.
XX
XX New zinc finger proteins, useful for modulating or regulating gene
XX expression and metabolic pathways in plants, e.g. for treating in the
XX plant cells a disorder that is associated with abnormal expression of the
XX target gene.
XX
XX Example 4; Page 148-149; 156pp; English.
XX
XX The patent discloses methods and compositions to modulate the expression
XX of a target gene in plant cells. The method involves providing plant
XX cells with a zinc finger protein (ZFP) which is capable of specifically
XX binding to a target nucleotide sequence or its complementary strand
XX within a target gene and allowing the ZFP binding to the target
XX nucleotide sequence, where the expression of the target gene in the plant
XX cells is modulated. The ZFP and fusions of the ZFP proteins are useful
XX for modulating or regulating gene expression and metabolic pathways in
XX plants. The ZFP, fusion proteins and methods are useful in plant and
XX agricultural technology. The method is useful particularly for treating a
XX disorder in the plant cells, where the disorder is associated with
XX abnormal expression of the target gene. The present DNA sequence is the
XX partial sequence of pMal- $\alpha$ 3 (APETALA3) and ZFP $\alpha$ 3 DNA
XX
XX Sequence 3300 BP; 853 A; 872 C; 876 G; 699 T; 0 U; 0 Other;
```

```
Query Match 97.8%; Score 3226.4; DB 4; Length 3300;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3254; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGTCGCAAAACCTTTTCGGGTATGCGATGATAGCCCGGAGAGA 60
```

Db 1 ||||| CCGACACCATCGAATGGTGC AAAACCTTTCGGGATATGGCATATGATAGCCCGGAGAGA 60
Qy 61 GTCAATTCCAGGGTGGTGAATGTGAACCCAGTAAACGTTATACGATGTTCGAGAGTATGCG 120
Db 61 GTCAATTCCAGGGTGGTGAATGTGAACCCAGTAAACGTTATACGATGTTCGAGAGTATGCG 120
Qy 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACCCAGGCGACGCAACGTTTCTGGGAAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACCCAGGCGACGCAACGTTTCTGGGAAAA 180
Qy 181 CCGCGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTTACATTCGCCACCGCTGGGCAC 240
Db 181 CCGCGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTTACATTCGCCACCGCTGGGCAC 240
Qy 241 AACAACTGGCGGCGAAAAAGTGTGCTGATTTGGGTTGGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCGAAAAAGTGTGCTGATTTGGGTTGGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 AGCGCGCTGCGCAAAATTTGTCGGCGGATTAATCTCGCGCCGATCAACTGGGTGCGAGCG 360
Db 301 AGCGCGCTGCGCAAAATTTGTCGGCGGATTAATCTCGCGCCGATCAACTGGGTGCGAGCG 360
Qy 361 TGGTGGTGTGATGGTGAACCAAGCGCGTCGAAGCCTGTAAAGCGGCGGTGCAACATC 420
Db 361 TGGTGGTGTGATGGTGAACCAAGCGCGTCGAAGCCTGTAAAGCGGCGGTGCAACATC 420
Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAACCTATCCGCTGGATGACCAAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAACCTATCCGCTGGATGACCAAGGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGGTTATTTCTGATGTCCTGACACAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGGTTATTTCTGATGTCCTGACACAGA 540
Qy 541 CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Qy 601 TGGTCCGATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGCGCCATTAAAGTCTGTCTCGG 660
Db 601 TGGTCCGATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGCGCCATTAAAGTCTGTCTCGG 660
Qy 661 CCGCTGTGCTGTGGCTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db 661 CCGCTGTGCTGTGGCTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Qy 721 CCGAACGGGAAGCGACTGGAGTGCATGTCGCTCGGTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CCGAACGGGAAGCGACTGGAGTGCATGTCGCTCGGTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATCGTTCCACTGCGATGCTGGTTGCCAAACGATCAGATGGCGCTGGGCGCAA 840
Db 781 ATGAGGGCATCGTTCCACTGCGATGCTGGTTGCCAAACGATCAGATGGCGCTGGGCGCAA 840
Qy 841 TCGCGCCATTACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
Db 841 TCGCGCCATTACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATCCCGCGGTTAACCCACCATCAAAACGAGATTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATCCCGCGGTTAACCCACCATCAAAACGAGATTTC 960
Qy 961 GCTGTGCGGGCAACACAGGTGGACGCTGTGTCGCAACTCTCTCAGGCGCCAGGCGTGA 1020
Db 961 GCTGTGCGGGCAACACAGGTGGACGCTGTGTCGCAACTCTCTCAGGCGCCAGGCGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGGCCGCTCTCATCTGGTGAAGAAAAACCAACCTGGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCGCTCTCATCTGGTGAAGAAAAACCAACCTGGGCGCCCAATA 1080
Qy 1081 CGGAAACCGCCTCTCCCGCGGCTTGGCCGATTCATTAATGACGTGGGACACAGAGTTT 1140
|||||

Db 1081 CGGAAACCGCCTCTCCCGCGGCTTGGCCGATTCATTAATGACGTGGGACACAGAGTTT 1140
Qy 1141 CCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Db 1141 CCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Qy 1201 GCACAAATTTCTATGTTTGACAGCTTATCATCGACTGCGGTCGACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATTTCTATGTTTGACAGCTTATCATCGACTGCGGTCGACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGACGCAATCGGAAGCTGTGGTATGGCTGTGAGGTCTGTAATCACTGCGTAAATTCG 1320
Db 1261 TCAGGACGCAATCGGAAGCTGTGGTATGGCTGTGAGGTCTGTAATCACTGCGTAAATTCG 1320
Qy 1321 TGTGCTCAAGCGGCACTCCCGTTCTGATTAATGTTTTCGCGCGACATCAATAACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCTGATTAATGTTTTCGCGCGACATCAATAACGGTT 1380
Qy 1381 CTGGCAATATTTCTGAAATGAGCTGTTCACAAATTAATCATCGGCTCGTATATGTGTGA 1440
Db 1381 CTGGCAATATTTCTGAAATGAGCTGTTCACAAATTAATCATCGGCTCGTATATGTGTGA 1440
Qy 1441 ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Qy 1501 GCATTTCCACCAACAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAAATCTGG 1560
Db 1501 GCATTTCCACCAACAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAAATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGT 1620
Qy 1621 ACCGGAATTAAGTCACCGTTGAGCATCCGGATAAATCTGGAAGAGAAAATTCACAGAGTT 1680
Db 1621 ACCGGAATTAAGTCACCGTTGAGCATCCGGATAAATCTGGAAGAGAAAATTCACAGAGTT 1680
Qy 1681 CCGGCAACTGGCGATGGCCCTGACATTAATCTTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CCGGCAACTGGCGATGGCCCTGACATTAATCTTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCTGTGGCTGAAATCAACCGGGAACAAAGGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTGGCTGAAATCAACCGGGAACAAAGGTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTAGTTTACAAACGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTAGTTTACAAACGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
Qy 1861 GAAGCGTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTTATAACAAAGATCTGCTGCCGNAACCCGCCAAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTCGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTCGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCCCTGCGCGCTGATTTGCTGCTGACGGGGGTTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCCCTGCGCGCTGATTTGCTGCTGACGGGGGTTTATCGGTTCAAG 2040
Qy 2041 TATGAAAAACGGCAAGTACGACATTAAGAAGCTGGGCGTGGATAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGGCAAGTACGACATTAAGAAGCTGGGCGTGGATAACGCTGGCGCGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGGTTGACCTGATTAAGAAACAAACACATGAATGCGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGGTTGACCTGATTAAGAAACAAACACATGAATGCGAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTCTTAATTAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTCTTAATTAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220

Qy	2221	GCATGGTCCAAATCATCGACACAGCAAAAGTGAATTAATGGTGTAAACGGTACTGCCGACCTTC	2280
Db	2221	GCATGGTCCAAATCATCGACACAGCAAAAGTGAATTAATGGTGTAAACGGTACTGCCGACCTTC	2380
Qy	2281	AAGGGTCAACCATCAAAACCGTTCGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT	2340
Db	2281	AAGGGTCAACCATCAAAACCGTTCGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT	2340
Qy	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACCTATCTGCTGACTGATGAAGGTCTG	2400
Db	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACCTATCTGCTGACTGATGAAGGTCTG	2400
Qy	2401	GAACCGGTTAATAAGACAAACCCCTGGGTGCGGTAGCGGTGAAGTCTTAACGAGGAAG	2460
Db	2401	GAACCGGTTAATAAGACAAACCCCTGGGTGCGGTAGCGGTGAAGTCTTAACGAGGAAG	2460
Qy	2461	TTGGCGAAAGATCCACGCTATTTCGCCGACCATGAAACGCCCAAGAAAGGTGAATCATG	2520
Db	2461	TTGGCGAAAGATCCACGCTATTTCGCCGACCATGAAACGCCCAAGAAAGGTGAATCATG	2520
Qy	2521	CCGAACATCCGCGAGATGTCCGCTTTCTGGTATGCCGTGCGTACTCGGTGATCAACGCC	2580
Db	2521	CCGAACATCCGCGAGATGTCCGCTTTCTGGTATGCCGTGCGTACTCGGTGATCAACGCC	2580
Qy	2581	GCCAGCGTGGTACATGTCGATGAAGCCCTGAAAGACGCCGACAGACTAATTCBAGCTCG	2640
Db	2581	GCCAGCGTGGTACATGTCGATGAAGCCCTGAAAGACGCCGACAGACTAATTCBAGCTCG	2640
Qy	2641	AACAAACAAACAAATAACAATAACAACAACTCGGGATCGAGGGAAGGATTCAGAAATTC	2700
Db	2641	AACAAACAAACAAATAACAATAACAACAACTCGGGATCGAGGGAAGGATTCAGAAATTC	2700
Qy	2701	GGATCTCTTCCTCTGTGGCCAGGCGGCGCTCGAGCCCGGGGAGAAAGCCCTATGCTTGT	2760
Db	2701	GGATCTCTTCCTCTGTGGCCAGGCGGCGCTCGAGCCCGGGGAGAAAGCCCTATGCTTGT	2760
Qy	2761	CCGGAATGGTAAGTCTCTCTCAGAGCTCTCACCTGGTGGCCACAGCGTACCCAC	2820
Db	2761	CCGGAATGGTAAGTCTCTCAGAGAGCTCTCTCGTGGCGCACAGCGTACCCAC	2820
Qy	2821	ACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTAGCCAGTCCAGCAAC	2880
Db	2821	ACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTAGCCAGTCCAGCAAC	2880
Qy	2881	CTGGTGGCCATCAACGCACTCATCTCTGGCGAGAGCCATCAAAATGTCAGAAATGTGGC	2940
Db	2881	CTGGTGGCCATCAACGCACTCATCTCTGGCGAGAGCCATCAAAATGTCAGAAATGTGGC	2940
Qy	2941	AAGTCTTTCTCTCGGTCTGACAACTCTCGTCCGGCACCAAGTACTCACACCGGGGAGAG	3000
Db	2941	AAGTCTTTAGCAGTCCAGCAACCTGGTGGCCACCAAGTACTCACACCGGGGAGAG	3000
Qy	3001	CCCTATGCTTGTCCGGAATGGTAAGTCTCTTTCAGCCGACAGGATAACCTGGTGCGCCAC	3060
Db	3001	CCCTATGCTTGTCCGGAATGGTAAGTCTCTTTCAGCACCAAGTGGCTCTTTGGTTAGACAC	3060
Qy	3061	CAGGTTACCCACAGCGGTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTTAGC	3120
Db	3061	CAGGTTACCCACAGCGGTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTTAGC	3120
Qy	3121	CAGCGCGGCCACTTGGCGAGCCATCAACGCACTCATCTCTGGCGAGAGCCATCAAAATGT	3180
Db	3121	CAGCGCGGCCACTTGGAAACCGCATCAACGCACTCATCTCTGGCGAGAGCCATCAAAATGT	3180
Qy	3181	CCAGAATGGGCAAGTCTTTCTCTCGGTGTGACAAATCTCGTCCGGCACCAAGTACTCAC	3240
Db	3181	CCAGAATGGGCAAGTCTTTCTCTCTCAACTTCAGGCAACTTTGGTCCGTCAACAACTACTCAC	3240
Qy	3241	ACCGGTAAAAAACTAGTGGCCAGGCGGCCAGTACCGGTACGACGTTCCGGACTACGCT	3300
Db	3241	ACCGGTAAAAAACTAGTGGCCAGGCGGCCAGTACCGGTACGACGTTCCGGACTACGCT	3300

RESULT 4	
ID	AAD11591 standard; DNA; 3300 BP.
XX	AC AAD11591;
XX	24-SEP-2001 (first entry)
DT	Partial sequence of pMal-m4 and ZFPm4 DNA.
DE	Zinc finger protein; ZFP; gene expression; metabolic pathway regulator
KW	modulation; plant technology; agriculture; ds.
KW	Unidentified.
XX	
OS	
XX	
FH	Key Location/Qualifiers
FT	CDS 2719..3270
FT	/tag= a
FT	/product= "ZFPm4 protein"
FT	/transl_except= [pos:3046..3048, aa:Ser]
FT	/note= "CDS does not include start and stop codon"
FT	/partial
FT	primer_bind 2740..2790
FT	/tag= b
FT	/bound_moiety= "F1-f2 primer"
FT	primer_bind 2770..2850
FT	/tag= c
FT	/bound_moiety= "F1-f1 primer"
FT	primer_bind 2824..2889
FT	/tag= d
FT	/bound_moiety= "F2-b primer"
FT	primer_bind 2867..2940
FT	/tag= e
FT	/bound_moiety= "F2-f primer"
FT	primer_bind 2916..2973
FT	/tag= f
FT	/bound_moiety= "F3-b1 primer"
FT	primer_bind 2953..3021
FT	/tag= g
FT	/bound_moiety= "F3-b2 primer"
FT	primer_bind 2992..3042
FT	/tag= h
FT	/bound_moiety= "F4-f2 primer"
FT	primer_bind 3022..3102
FT	/tag= i
FT	/bound_moiety= "F4-f1 primer"
FT	primer_bind 3076..3141
FT	/tag= j
FT	/bound_moiety= "F5-b primer"
FT	primer_bind 3119..3192
FT	/tag= k
FT	/bound_moiety= "F5-f primer"
FT	primer_bind 3168..3225
FT	/tag= l
FT	/bound_moiety= "F6-b1 primer"
FT	primer_bind 3205..3273
FT	/tag= m
FT	/bound_moiety= "F6-b2 primer"
XX	
PN	WO200152620-A2.
PD	26-JUL-2001.
XX	
PF	19-JAN-2001; 2001WO-US001817.
XX	
PR	21-JAN-2000; 2000US-0177468P.
PR	21-JUL-2000; 2000US-00620897.
XX	
PA	(SCRI) SCRIPPS RES INST.
PA	(SYGN) SYNGENTA AGRIC DISCOVERY INC.
XX	
PI	Barbas CF, Stece JT, Guan X, Dalmia B:

XX WPI; 2001-465325/50.
DR P-PSDB; AAE06004.
XX
PT New zinc finger proteins, useful for modulating or regulating gene
PT expression and metabolic pathways in plants, e.g. for treating in the
PT plant cells a disorder that is associated with abnormal expression of the
PT target gene.
XX
PS Example 4; Page 145-147; 156pp; English.
XX
CC The patent discloses methods and compositions to modulate the expression
CC of a target gene in plant cells. The method involves providing plant
CC cells with a zinc finger protein (ZFP) which is capable of specifically
CC binding to a target nucleotide sequence or its complementary strand
CC within a target gene and allowing the ZFP binding to the target
CC nucleotide sequence, where the expression of the target gene in the plant
CC cells is modulated. The ZFP and fusions of the ZFP proteins are useful
CC for modulating or regulating gene expression and metabolic pathways in
CC plants. The ZFP, fusion proteins and methods are useful in plant and
CC agricultural technology. The method is useful particularly for treating a
CC disorder in the plant cells, where the disorder is associated with
CC abnormal expression of the target gene. The present DNA sequence is the
CC partial sequence of pMal-m4 and ZFPm4 DNA
XX
SQ Sequence 3300 BP; 849 A; 865 C; 880 G; 706 T; 0 U; 0 Other;
Query Match 97.4%; Score 3215.2; DB 4; Length 3300;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 3247; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
QY 1 CCGACACCATCAATGGTGCAGAAACCTTTCCGGGTATGCGATAGAGCCCGGAAGAGA 60
DB 1 CCGACACCATCAATGGTGCAGAAACCTTTCCGGGTATGCGATAGAGCCCGGAAGAGA 60
QY 61 GTCATTCAGGTGGTGAATGTGAACAGTAACGTATACCATCTCCGAGAGTATCGG 120
DB 61 GTCATTCAGGTGGTGAATGTGAACAGTAACGTATACCATCTCCGAGAGTATCGG 120
QY 121 GTGCTCTTATCAGACCGGTTTCCCGCGTGGTGAACAGGCCACGACGTTTCTGCGAAAA 180
DB 121 GTGCTCTTATCAGACCGGTTTCCCGCGTGGTGAACAGGCCACGACGTTTCTGCGAAAA 180
QY 181 CCGGGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTTACATTCCTCCAAACCGCGTGGCAC 240
DB 181 CCGGGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTTACATTCCTCCAAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAACAGTCTGTTGCTGATTTGGCGTTTGGCCACCTCCAGTCTGGGCCCTGC 300
DB 241 AACAACTGGCGGGCAACAGTCTGTTGCTGATTTGGCGTTTGGCCACCTCCAGTCTGGGCCCTGC 300
QY 301 ACGCGCGCTGCGCAAAATTTGTCGGCGGATTAATACTCTCGCGCGATCAACTGGGTGGCAGCG 360
DB 301 ACGCGCGCTGCGCAAAATTTGTCGGCGGATTAATACTCTCGCGCGATCAACTGGGTGGCAGCG 360
QY 361 TGGTGGTTCGATGTAGAACGAAGCGCGTGAAGCCCTGTAAAGCGCGGTGCAACAATC 420
DB 361 TGGTGGTTCGATGTAGAACGAAGCGCGTGAAGCCCTGTAAAGCGCGGTGCAACAATC 420
QY 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCAATTAATCTCGCTGGATGACAGGATGCCA 480
DB 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCAATTAATCTCGCTGGATGACAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTGTACCCAGA 540
DB 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTGTACCCAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
QY 601 TGGTGCATTTGGGTACACGAAATTCGCGCTGTTAGCGGGGCCATTAAGTTCTGTCTCGG 660
DB 601 TGGTGCATTTGGGTACACGAAATTCGCGCTGTTAGCGGGGCCATTAAGTTCTGTCTCGG

DB 601 TGGTGCATTTGGGTACACGAAATTCGCGCTGTTAGCGGGGCCATTAAGTTCTGTCTCGG 660
QY 661 CCGGTCTCGCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CCGGTCTCGCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTCAACAAACCAATGCAAAATCTCTGA 780
DB 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTCAACAAACCAATGCAAAATCTCTGA 780
QY 781 ATGAGGGCATCGTTCCTCCACTGCGATGCTGGTTGCTGCAACGATCAGATGGCGCTGGCGGCAA 840
DB 781 ATGAGGGCATCGTTCCTCCACTGCGATGCTGGTTGCTGCAACGATCAGATGGCGCTGGCGGCAA 840
QY 841 TGGCGGCCATTAACGAGTCCCGGCTGCGGTTGGTGGCGGATATCTCGGTAGTGGGATAG 900
DB 841 TGGCGGCCATTAACGAGTCCCGGCTGCGGTTGGTGGCGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAAAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAAAACACGAGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
DB 961 GCCTGCTGGGCAAAACACGAGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAAAAACCCCTGGCGCCCAATA 1080
DB 1021 AGGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAAAAACCCCTGGCGCCCAATA 1080
QY 1081 CGAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGCTGGCACGACAGGTTT 1140
DB 1081 CGAAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGCTGGCACGACAGGTTT 1140
QY 1141 CCGACTGCGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCTACTCATTAG 1200
DB 1141 CCGACTGCGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCTACTCATTAG 1200
QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGCGTGCACCAATGTTCTGGCG 1260
DB 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGCGTGCACCAATGTTCTGGCG 1260
QY 1261 TCAGGACCGCATCGGAAGCTGTGATGCTGTGCGAGTTCGTAATCATCTGATTAATTCG 1320
DB 1261 TCAGGACCGCATCGGAAGCTGTGATGCTGTGCGAGTTCGTAATCATCTGATTAATTCG 1320
QY 1321 TGTGCTCTCAAGCGGCACTCCCGTTCTGGATATGTTTTTTCGCGCGACATCAATAACGGTT 1380
DB 1321 TGTGCTCTCAAGCGGCACTCCCGTTCTGGATATGTTTTTTCGCGCGACATCAATAACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
DB 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
QY 1441 ATTGTGAGCGGATTAACAATTTTCAACAGGAAACAGCCAGTCCGTTAGGTGTTTTCACGA 1500
DB 1441 ATTGTGAGCGGATTAACAATTTTCAACAGGAAACAGCCAGTCCGTTAGGTGTTTTCACGA 1500
QY 1501 GCACCTTCAACCAACAGGACCATAGATTATGAAAACTGAAGAGGTTAACTGGTAAATCTCG 1560
DB 1501 GCACCTTCAACCAACAGGACCATAGATTATGAAAACTGAAGAGGTTAACTGGTAAATCTCG 1560
QY 1561 ATTAACGGCGATTAAGGCTATAACCGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTAACGGCGATTAAGGCTATAACCGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTTCCACAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTTCCACAGGTT 1680
QY 1681 GGGGCAATCGGCGATGGCCCTGACATTAATCTCTGGGCACACGCGCTTGTGGCTAC 1740
DB 1681 GGGGCAATCGGCGATGGCCCTGACATTAATCTCTGGGCACACGCGCTTGTGGCTAC 1740

QY	1741	GCTCAATCTGGCTCTGTGGCTGAAATCACCCCGGCAAAAGCGTTCCAGGACAAGCTGTAT	1800
Db	1741	GCTCAATCTGGCTCTGTGGCTGAAATCACCCCGGCAAAAGCGTTCCAGGACAAGCTGTAT	1800
QY	1801	CGTTTACCTGGATGCGGTACGTTTAAACGGCAAGCTGATTTGCTTACCCGATCGCTGT	1860
Db	1801	CGTTTACCTGGATGCGGTACGTTTAAACGGCAAGCTGATTTGCTTACCCGATCGCTGT	1860
QY	1861	GAAGGTTTATCGCTGATTTAAACAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA	1920
Db	1861	GAAGGTTTATCGCTGATTTAAACAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA	1920
QY	1921	GAGATCCCGGGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGGGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACGCTATTACCTGGCGCTGATTTGCTGTCGAGCGGGGTTATGCGTTCAAG	2040
Db	1981	CTGCAAGAACGCTATTACCTGGCGCTGATTTGCTGTCGAGCGGGGTTATGCGTTCAAG	2040
QY	2041	TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATACGCTGGCGCGAAGCG	2100
Db	2041	TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATACGCTGGCGCGAAGCG	2100
QY	2101	GGTCTGACCTTCTGCTGTTGACCTGATTTAAAAACAAACACATGAATGCAAGACCGGATTAC	2160
Db	2101	GGTCTGACCTTCTGCTGTTGACCTGATTTAAAAACAAACACATGAATGCAAGACCGGATTAC	2160
QY	2161	TCCATCGAAGAGCTGCTTTTAAAGGCGAAACAGCGATGACATCAACGGCCCGTGG	2220
Db	2161	TCCATCGAAGAGCTGCTTTTAAAGGCGAAACAGCGATGACATCAACGGCCCGTGG	2220
QY	2221	GCATGGTCAACATCGACACAGCAAAAGTGAATTTATGGTGAACGGTACTGCCACCTTC	2280
Db	2221	GCATGGTCAACATCGACACAGCAAAAGTGAATTTATGGTGAACGGTACTGCCACCTTC	2280
QY	2281	AAGGGTCAACCATCAAAACCGTTCTGGTGGCTGCTGAGCGCAGGTATTAACGCCGCCAGT	2340
Db	2281	AAGGGTCAACCATCAAAACCGTTCTGGTGGCTGCTGAGCGCAGGTATTAACGCCGCCAGT	2340
QY	2341	CCGAACAAAGAGCTGGCAAAAGAGTCTCTGAAAACCTATCTGCTGACTGATGAAGGCTG	2400
Db	2341	CCGAACAAAGAGCTGGCAAAAGAGTCTCTGAAAACCTATCTGCTGACTGATGAAGGCTG	2400
QY	2401	GAAGCGGTTAATAAGACAAACCGCTGGTGGCTGCTGAGCGCAGGTATTAACGCCGCCAG	2460
Db	2401	GAAGCGGTTAATAAGACAAACCGCTGGTGGCTGCTGAGCGCAGGTATTAACGCCGCCAG	2460
QY	2461	TTGGCGAAAGATCCAGCTATTTGCCGCCACCATGGAACACGCCCAAGAGGTGAATCATG	2520
Db	2461	TTGGCGAAAGATCCAGCTATTTGCCGCCACCATGGAACACGCCCAAGAGGTGAATCATG	2520
QY	2521	CCGAACATCCGCAAGATGCTGGCTTTCTGGTATGCTGCTGCTGCTGCTGCTGCTGCTG	2580
Db	2521	CCGAACATCCGCAAGATGCTGGCTTTCTGGTATGCTGCTGCTGCTGCTGCTGCTGCTG	2580
QY	2581	GCCAGCGGTCTGAGCTGCTGAGTGAAGCGCTGAAAGACGCGAGACTAATTCGAGCTCG	2640
Db	2581	GCCAGCGGTCTGAGCTGCTGAGTGAAGCGCTGAAAGACGCGAGACTAATTCGAGCTCG	2640
QY	2641	AACAAACAAACAAATACAAACAACTCGGGATCGAGGGAAGGATTTTTCAGAAATTC	2700
Db	2641	AACAAACAAACAAATACAAACAACTCGGGATCGAGGGAAGGATTTTTCAGAAATTC	2700
QY	2701	GGATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2760
Db	2701	GGATCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2760
QY	2761	CCGGAATGTGGTAAAGTCTTCTCTCAGAGCTCTCACCTGGTGGCCACACAGCTATCCAC	2820
Db	2761	CCGGAATGTGGTAAAGTCTTCTCTCAGAGCTCTCACCTGGTGGCCACACAGCTATCCAC	2820

QY	2821	ACGGGTGAAAACCGTATTAATGCCAGAGTGGCGGAAAATCTTTTAGCAGTCCAGCAAC	2880
Db	2821	ACGGGTGAAAACCGTATTAATGCCAGAGTGGCGGAAAATCTTTTAGCAGCAGCAGC	2880
QY	2881	CTGGTGGCCCATCAACCGCACTCATCTGCGGAGAAAGCATACAATGTCCAGAATGTGGC	2940
Db	2881	CTGGTGGCCCATCAACCGCACTCATCTGCGGAGAAAGCATACAATGTCCAGAATGTGGC	2940
QY	2941	AAGTCTTTCTCGGTCTGCAAACTCTCGTCCGGCACCAACGTACTCACACCGGGGAGAAG	3000
Db	2941	AAGTCTTTCTCGGTCTGCAAACTCTCGTCCGGCACCAACGTACTCACACCGGGGAGAAG	3000
QY	3001	CCCTATCTGCTGCTGGGAATGTGGTAACTCTCTAGCCGCGAGCATACCTGGTGGCCAC	3060
Db	3001	CCCTATCTGCTGCTGGGAATGTGGTAACTCTCTAGCCGCGAGCATACCTGGTGGCCAC	3060
QY	3061	CAGGTTACCCACACCGGTGAAAACCGTATTAATGCCAGAGTGGCGGAAAATCTTTTAGC	3120
Db	3061	CAGGTTACCCACACCGGTGAAAACCGTATTAATGCCAGAGTGGCGGAAAATCTTTTAGC	3120
QY	3121	CAGCCCGGCGACCTGGCCAGCCATCAACGCACTCATCTGCGGAGAAAGCATACAATGT	3180
Db	3121	CAGCCCGGCGACCTGGCCAGCCATCAACGCACTCATCTGCGGAGAAAGCATACAATGT	3180
QY	3181	CCAGAAATGTGCGAAGTCTTTCTCTCGGTCTGCAAACTCTGCTGGGACCAACGTACTAC	3240
Db	3181	CCAGAAATGTGCGAAGTCTTTCTCTCGGTCTGCAAACTCTGCTGGGACCAACGTACTAC	3240
QY	3241	ACCGGTAAAAAAACTAGTGGCCAGGCGCGCAGTACCCGTAACGCTTCCGGACTACGCT	3300
Db	3241	ACCGGTAAAAAAACTAGTGGCCAGGCGCGCAGTACCCGTAACGCTTCCGGACTACGCT	3300

RESULT 5
AAD11590
ID AAD11590 standard; DNA; 3300 BP.
XX
AC AAD11590;
XX
DT 24-SEP-2001 (first entry)
XX
DE Partial sequence of pMal-m3 and ZFPm3 DNA.
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
KW modulation; plant technology; agriculture; ds.
XX
OS Unidentified.
XX
Key Location/Qualifiers
CDS 2719..3270
FT /*tag= a
FT /product= "ZFPm3 protein"
FT /note= "CDS does not include start and stop codon"
FT /partial
FT primer_bind 2740..2790
FT /*tag= b
FT /bound_moiety= "F1-f2 primer"
FT primer_bind 2770..2850
FT /*tag= c
FT /bound_moiety= "F1-f1 primer"
FT primer_bind 2824..2889
FT /*tag= d
FT /bound_moiety= "F2-b primer"
FT primer_bind 2867..2940
FT /*tag= e
FT /bound_moiety= "F2-f primer"
FT primer_bind 2916..2973
FT /*tag= f
FT /bound_moiety= "F3-b1 primer"
FT primer_bind 2953..3021
FT /*tag= g
FT /bound_moiety= "F3-b2 primer"
FT primer_bind 2992..3042

```
FT      /*tag= h
FT      /bound_moiety= "F4-f2 primer"
FT      3022..3102
FT      /*tag= i
FT      /bound_moiety= "F4-f1 primer"
FT      3076..3141
FT      /*tag= j
FT      /bound_moiety= "F5-b primer"
FT      3119..3192
FT      /*tag= k
FT      /bound_moiety= "F5-f primer"
FT      3168..3225
FT      /*tag= l
FT      /bound_moiety= "F6-b1 primer"
FT      3205..3273
FT      /*tag= m
FT      /bound_moiety= "F6-b2 primer"
XX
XX      WO200152620-A2.
XX
XX      26-JUL-2001.
XX
XX      19-JAN-2001; 2001WO-US001817.
XX
XX      21-JAN-2000; 2000US-0177468P.
XX      21-JUL-2000; 2000US-00620897.
XX
XX      (SCRI ) SCRIPPS RES INST.
XX      (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
XX      Barbas CF, Stege JT, Guan X, Dalmia B;
XX      WPI; 2001-465325/50.
XX      P-PSDB; AA06003.
XX
XX      New zinc finger proteins, useful for modulating or regulating gene
XX      expression and metabolic pathways in plants, e.g. for treating in the
XX      plant cells a disorder that is associated with abnormal expression of the
XX      target gene.
XX
XX      Example 4; Page 143-145; 156pp; English.
XX
XX      The patent discloses methods and compositions to modulate the expression
XX      of a target gene in plant cells. The method involves providing plant
XX      cells with a zinc finger protein (ZFP) which is capable of specifically
XX      binding to a target nucleotide sequence or its complementary strand
XX      within a target gene and allowing the ZFP binding to the target
XX      nucleotide sequence, where the expression of the target gene in the plant
XX      cells is modulated. The ZFP and fusions of the ZFP proteins are useful
XX      for modulating or regulating gene expression and metabolic pathways in
XX      plants. The ZFP, fusion proteins and methods are useful in plant and
XX      agricultural technology. The method is useful particularly for treating a
XX      disorder in the plant cells, where the disorder is associated with
XX      abnormal expression of the target gene. The present DNA sequence is the
XX      Partial sequence of pMal-m3 and ZFPm3 DNA
XX
XX      Sequence 3300 BP; 845 A; 878 C; 877 G; 700 T; 0 U; 0 Other;
XX
XX      Query Match      97.2%; Score 3207.2; DB 4; Length 3300;
XX      Best Local Similarity 98.2%; Pred. No. 0;
XX      Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
XX
XX      1  CGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGGCCCGCGGAAGAGA 60
XX      Db      |||
XX      1  CGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGGCCCGCGGAAGAGA 60
XX
XX      61  GTCAATTACAGGTGGTGAATGTGAACCAAGTAAACGTTATACGATTCGCAGAGTATGCG 120
XX      Db      |||
XX      61  GTCAATTACAGGTGGTGAATGTGAACCAAGTAAACGTTATACGATTCGCAGAGTATGCG 120
XX
XX      121  GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGCCAGCCACGTTTCTGCGAAA 180
XX      Db      |||
XX      121  GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGCCAGCCACGTTTCTGCGAAA 180
```

```
Qy      181  CGCGGAAAAAGTGGAAAGCGCGGATGGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
Dy      |||
Dy      181  CGCGGAAAAAGTGGAAAGCGCGGATGGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
Qy      241  AACAACTGGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300
Dy      |||
Dy      241  AACAACTGGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300
Qy      301  ACGCGCGTCGCAAAATTTGTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCGACGG 360
Dy      |||
Dy      301  ACGCGCGTCGCAAAATTTGTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCGACGG 360
Qy      361  TGGTGGTGTGATGGTAGAACGAGCGCGCTGGAAGCTGTAAAGCGCGGTGCACAATC 420
Dy      |||
Dy      361  TGGTGGTGTGATGGTAGAACGAGCGCGCTGGAAGCTGTAAAGCGCGGTGCACAATC 420
Qy      421  TTCTCGCGCAACGCTCAGTCGGGCTGATCAATTAATCTCCGCTGGATCACCAGGATGCCA 480
Dy      |||
Dy      421  TTCTCGCGCAACGCTCAGTCGGGCTGATCAATTAATCTCCGCTGGATCACCAGGATGCCA 480
Qy      481  TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 540
Dy      |||
Dy      481  TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 540
Qy      541  CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGCTGGCGGTGGAGCATC 600
Dy      |||
Dy      541  CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGCTGGCGGTGGAGCATC 600
Qy      601  TGGTGGCATTCGGCTCACCAGCAAAATCGGCTGTAGCGGGGCCATTAGTTCTGTCGCG 660
Dy      |||
Dy      601  TGGTGGCATTCGGCTCACCAGCAAAATCGGCTGTAGCGGGGCCATTAGTTCTGTCGCG 660
Qy      661  CGCGTCTCGCTCGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Dy      |||
Dy      661  CGCGTCTCGCTCGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy      721  CGGAAACGGGAAAGGCGATGGAGTGCCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Dy      |||
Dy      721  CGGAAACGGGAAAGGCGATGGAGTGCCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Qy      781  ATGAGGGCATGTTCCCACTGGGATGCTGGTTCGCAACGATCAGATGCGGTGGCGGCA 840
Dy      |||
Dy      781  ATGAGGGCATGTTCCCACTGGGATGCTGGTTCGCAACGATCAGATGCGGTGGCGGCA 840
Qy      841  TGGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAGC 900
Dy      |||
Dy      841  TGGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAGC 900
Qy      901  ACGATACCGAAGACAGCTCATGTTATATPCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
Dy      |||
Dy      901  ACGATACCGAAGACAGCTCATGTTATATPCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
Qy      961  GCCTGTGGGGCAACACGAGCTGGACCGCTTGTGCAACTCTCTCAGGCCCGCGGTGA 1020
Dy      |||
Dy      961  GCCTGTGGGGCAACACGAGCTGGACCGCTTGTGCAACTCTCTCAGGCCCGCGGTGA 1020
Qy      1021  AGGGCAATCAGCTGTTGCCCTCTCACTGGTGAAGAAACCAACCCCTGGCGGCCAATA 1080
Dy      |||
Dy      1021  AGGGCAATCAGCTGTTGCCCTCTCACTGGTGAAGAAACCAACCCCTGGCGGCCAATA 1080
Qy      1081  CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCATTAATGCAAGTGGCACGACAGGTTT 1140
Dy      |||
Dy      1081  CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCATTAATGCAAGTGGCACGACAGGTTT 1140
Qy      1141  CCCGACTGGAAGCGGCGAGTGAGGCAACCGAATTAATGAGTGGTACTCACTCATTAG 1200
Dy      |||
Dy      1141  CCCGACTGGAAGCGGCGAGTGAGGCAACCGAATTAATGAGTGGTACTCACTCATTAG 1200
Qy      1201  GCACAATTCATGTTTTCACAGCTTATCATGACCTGACGCGGTGACCAATGCTTCTGGCG 1260
Dy      |||
Dy      1201  GCACAATTCATGTTTTCACAGCTTATCATGACCTGACGCGGTGACCAATGCTTCTGGCG 1260
```


QY	1261	TCAGGAGCCATCGAAGCTGTGGTATAGCTGTGAGGTCTGTAATCACTGCTAATTCG	1320
Db	1261	TCAGGAGCCATCGAAGCTGTGGTATAGCTGTGAGGTCTGTAATCACTGCTAATTCG	1320
QY	1321	TGTCGCTCAAGCGCACCTCCGTTCTGGATAATGTTTTTGGCGCGACATCAATCAAGGTT	1380
Db	1321	TGTCGCTCAAGCGCACCTCCGTTCTGGATAATGTTTTTGGCGCGACATCAATCAAGGTT	1380
QY	1381	CTGGCAATATTTCTGAATAGCTGTGCACATTAATCATCGGCTCGTAAATGTTGCGA	1440
Db	1381	CTGGCAATATTTCTGAATAGCTGTGCACATTAATCATCGGCTCGTAAATGTTGCGA	1440
QY	1441	ATTGTAGCGGATAACAATTTACACAGAAACAGCCAGTCCGTTTAGTGTTTTCAAGA	1500
Db	1441	ATTGTAGCGGATAACAATTTACACAGAAACAGCCAGTCCGTTTAGTGTTTTCAAGA	1500
QY	1501	GCACCTTACCACAGGACCATAGATTATGAAACTGAAGCTGAAGGTAAACTGGTAACTCGG	1560
Db	1501	GCACCTTACCACAGGACCATAGATTATGAAACTGAAGGTAAACTGGTAACTCGG	1560
QY	1561	ATTAAACGCGATAAAGGCTATAACGCTCTCGTGAAGTCGGTAAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAAACGCGATAAAGGCTATAACGCTCTCGTGAAGTCGGTAAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAAGTCACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCGCCACAGTT	1680
Db	1621	ACCGGAATTAAGTCACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCGCCACAGTT	1680
QY	1681	CGGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGCGTTTGGTGCTAC	1740
Db	1681	CGGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGCGTTTGGTGCTAC	1740
QY	1741	GCTCAATCGCCCTGTGGCTGAATATCAACCCGACAAAGGTTCCAGGACAAAGCTGTAT	1800
Db	1741	GCTCAATCGCCCTGTGGCTGAATATCAACCCGACAAAGGTTCCAGGACAAAGCTGTAT	1800
QY	1801	CCGTTTACCTGGATGCGGTACGTTACAAACGGAAGCTGATGCTTACCGGATCGCTGTT	1860
Db	1801	CCGTTTACCTGGATGCGGTACGTTACAAACGGAAGCTGATGCTTACCGGATCGCTGTT	1860
QY	1861	GAAGGTTATCGCTGATTATTAACAAGATCTGTCGCGAACCCGCAAAACCTCGGAA	1920
Db	1861	GAAGGTTATCGCTGATTATTAACAAGATCTGTCGCGAACCCGCAAAACCTCGGAA	1920
QY	1921	GAGATCCGCGCGTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCGCGCGTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTCGCGCTGATGCTGCTGACGCGGCTTATGCTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTCACTCGCGCGCTGATGCTGCTGACGCGGCTTATGCTTCAAG	2040
QY	2041	TATGAAACCGCAAGTACGACATTAAGAAGCTGGGCGTGGATACGCTGGCGGAAAGCG	2100
Db	2041	TATGAAACCGCAAGTACGACATTAAGAAGCTGGGCGTGGATACGCTGGCGGAAAGCG	2100
QY	2101	GGTCTGACCTTCTGGTTGACCTGATTAAACAAACACATGAATGAGACACCGATTAC	2160
Db	2101	GGTCTGACCTTCTGGTTGACCTGATTAAACAAACACATGAATGAGACACCGATTAC	2160
QY	2161	TCCATCGCAGAGCTGCTTTTAAAGGCGAAACAGGATGACCATCAACGCGCGCTGG	2220
Db	2161	TCCATCGCAGAGCTGCTTTTAAAGGCGAAACAGGATGACCATCAACGCGCGCTGG	2220
QY	2221	GCATGGTCCAACTCGACACAGCAAAAGTGAATTTATGGTGAACGCTTACCGGACCTTC	2280
Db	2221	GCATGGTCCAACTCGACACAGCAAAAGTGAATTTATGGTGAACGCTTACCGGACCTTC	2280
QY	2281	AAGGTCACCACTCAACCGTTCTGGGCTGCTGAGCGCAGGTATTAACGCGCGCAGT	2340
Db	2281	AAGGTCACCACTCAACCGTTCTGGGCTGCTGAGCGCAGGTATTAACGCGCGCAGT	2340
QY	2341	CCGAAACAAAGAGCTGGCAAAAGATTCTCTCGAAAACTATCTGCTGACTGATGAAGGCTG	2400

Db	2341	CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGCTCTG	2400
QY	2401	GAAGCGGTTAATAAAGACAAACCCGCTGGGTGCGGTAGCGTGAAGTCTTTACGAGGAAGAG	2460
Db	2401	GAAGCGGTTAATAAAGACAAACCCGCTGGGTGCGGTAGCGTGAAGTCTTTACGAGGAAGAG	2460
QY	2461	TTGGCGAAAGATCCACGTTATTTGCCGCCACCATGGAAGAAACGCCAGAAAGGTGAATCATG	2520
Db	2461	TTGGCGAAAGATCCACGTTATTTGCCGCCACCATGGAAGAAACGCCAGAAAGGTGAATCATG	2520
QY	2521	CCGAAACATCCGCGAGATGTCGGCTTTCTGTTATGTCGCTGCTGCTGCTGCTGCTGCTG	2580
Db	2521	CCGAAACATCCGCGAGATGTCGGCTTTCTGTTATGTCGCTGCTGCTGCTGCTGCTGCTG	2580
QY	2581	GCACGCGCTCGTCAAGTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
Db	2581	GCACGCGCTCGTCAAGTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
QY	2641	AACAAACAAACAAATTAACAACCTCGGGATCGAGGAGAGGATTCAGAAATTC	2700
Db	2641	AACAAACAAACAAATTAACAACCTCGGGATCGAGGAGAGGATTCAGAAATTC	2700
QY	2701	GGATCTCTTCTCTGTCGCGCGGCGCTCGAGCCCGGGGAGAGCCCTATGCTTGT	2760
Db	2701	GGATCTCTTCTCTGTCGCGCGGCGCTCGAGCCCGGGGAGAGCCCTATGCTTGT	2760
QY	2761	CCGGAATGTTGTAAGTCTTCTCTGAGGCTCTCACTGTCGCTGCTGCTGCTGCTGCTGCTG	2820
Db	2761	CCGGAATGTTGTAAGTCTTCTCTGAGGCTCTCACTGTCGCTGCTGCTGCTGCTGCTGCTG	2820
QY	2821	ACGGGTGAAAAACCGGTATAATGCCAGAGTGGCGGAAATCTTTTAGCCAGTCCAGCAAC	2880
Db	2821	ACGGGTGAAAAACCGGTATAATGCCAGAGTGGCGGAAATCTTTTAGCCAGTCCAGCAAC	2880
QY	2881	CTGCTGCGCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCAGAAATGTGGC	2940
Db	2881	CTGCTGCGCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCAGAAATGTGGC	2940
QY	2941	AAGTCTTCTCTGCTGCTGCAATCTGTCGCGGACCAAGCTACTCACACCGGGGAGAAG	3000
Db	2941	AAGTCTTCTCTGCTGCTGCAATCTGTCGCGGACCAAGCTACTCACACCGGGGAGAAG	3000
QY	3001	CCCTATGCTTGTCCGGAATGTGGTAAAGTCTTCTCAGCGCGAGGATAACCTGTGCGCCAC	3060
Db	3001	CCCTATGCTTGTCCGGAATGTGGTAAAGTCTTCTCAGCGCGAGGATAACCTGTGCGCCAC	3060
QY	3061	CAGCGTACCCACACCGGGTGAATAACCGTATAATGCGGAGTGGCGGAAATCTTTTAGT	3120
Db	3061	CAGCGTACCCACACCGGGTGAATAACCGTATAATGCGGAGTGGCGGAAATCTTTTAGT	3120
QY	3121	CAGCGCGCGCACCTGCGCCAGCCATCAACGCACTCATCTGGCGAGAGCCATCAAAATGT	3180
Db	3121	CAGCGCGCGCACCTGCGCCAGCCATCAACGCACTCATCTGGCGAGAGCCATCAAAATGT	3180
QY	3181	CCAGAAATGTGGCAAGTCTTTCTCTGCTGCTGCAATCTGTCGCGGACCAACGCTACTCAC	3240
Db	3181	CCAGAAATGTGGCAAGTCTTTCTCTGCTGCTGCAATCTGTCGCGGACCAACGCTACTCAC	3240
QY	3241	ACCGGTAAAAAACTAGTGGCGGCGGCGGCTAGTACCCGTAACGCTTCCGGACTACGCT	3300
Db	3241	ACCGGTAAAAAACTAGTGGCGGCGGCGGCTAGTACCCGTAACGCTTCCGGACTACGCT	3300

RESULT 6

AAH25681

ID AAH25681 standard; DNA; 8101 BP.

XX

AC AAH25681;

XX

DT 05-SBP-2001 (first entry)

XX

DE Nucleotide sequence of human Cdc25C phosphatase DNA..

XX Fusion protein; maltose binding protein; MBP; Cdc25C phosphatase;
KW Cdc25B1; Cdc25B2; Cdc25B3; cyclin-dependent kinase; CDK; cell division;
KW ss.
XX
XX Homo sapiens.
OS
XX WO200144467-A2.
PN
XX 21-JUN-2001.
PD
XX
PF 13-DEC-2000; 2000WO-PR003496.
XX
XX 14-DEC-1999; 99FR-00015722.
PR
XX 30-MAY-2000; 2000FR-00006883.
PR
XX 21-SEP-2000; 2000FR-00012008.
PR
XX (SCRC) SCRAS SOC CONSEILS RECH & APPL SCI.
PA
XX Goubin-Gramatica F, Ducommun B, Prevost G;
PI
XX WPI; 2001-398152/42.
DR
XX
XX New fusion protein of maltose-binding protein and Cdc phosphatase, useful
PT for identifying phosphatase modulators for regulating the cell cycle.
XX
XX Claim 7; Page 9-13; 56pp; French.
XX
XX The present sequence encodes human Cdc25C phosphatase, and was expressed
CC in Escherichia coli. It is used to make fusion proteins with Escherichia
CC coli maltose binding protein (MBP). The specification describes fusion
CC proteins of MBP and one of Cdc 25B1, 25B2, 25B3 and 25C. Cdc phosphatases
CC are involved in activation of cyclin-dependent kinases (CDK) that control
CC cell division. The fusion proteins are used to identify modulators of the
CC specified human Cdc phosphatases, potentially useful for regulation of
CC cell division. They are also for studying physiological or
CC physiopathological activities of Cdc phosphatase
XX
SQ Sequence 8101 BP; 2127 A; 1987 C; 2121 G; 1866 T; 0 U; 0 Other;

Query Match 82.0%; Score 2705.8; DB 4; Length 8101;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CGGACACCATCAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
DB 1 CGGACACCATCAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTACAGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGAGAGTATGCCG 120
DB 61 GTCAATTACAGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGAGAGTATGCCG 120
QY 121 GTGTCTTATCAGACCGTTTCCGGGTGGTGAACAGGCCAGCCACGCTTCTGCGAATA 180
DB 121 GTGTCTTATCAGACCGTTTCCGGGTGGTGAACAGGCCAGCCACGCTTCTGCGAATA 180
QY 181 CGCGGGAAAAAGTGAAGCGCGGATCGCGAGCTGAATTTACATTCGCAACCGGTGGCAC 240
DB 181 CGCGGGAAAAAGTGAAGCGCGGATCGCGAGCTGAATTTACATTCGCAACCGGTGGCAC 240
QY 241 AACAACTGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 AGCGCGGTTCGCAAAATGTTCGGGGGATTAATCTCGCGCCGATCAACTGGGTGGCAGCG 360
DB 301 AGCGCGGTTCGCAAAATGTTCGGGGGATTAATCTCGCGCCGATCAACTGGGTGGCAGCG 360
QY 361 TGGTGTGTTCGATGTGAGAACGAGCGGTTCGAGCCCTGTAAGCGCGGTGCAAAATC 420
DB 361 TGGTGTGTTCGATGTGAGAACGAGCGGTTCGAGCCCTGTAAGCGCGGTGCAAAATC 420
QY 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACCTATCCGCTGGATGACAGGATGCCA 480

Db 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATCTCCGCTGGATGACACAGGATGCCA 480
QY 481 TTGCTGTGGAAGAGCTGCTGCACATAATGTTCCGGCGCTTATTTCTTGATGCTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGAGCTGCTGCACATAATGTTCCGGCGGTATTTCTTGTGATGCTCTGACCCAGA 540
QY 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGACGCTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGACGCTACGCGACTGGGCGTGGAGCATC 600
QY 601 TGGTCGCATTCGGTCACCAAGCAAAATTCGCGCTGTGTAGCGGGCCCAATAAGTTCTGCTCGG 660
Db 601 TGGTCGCATTCGGTCACCAAGCAAAATTCGCGCTGTGTAGCGGGCCCAATAAGTTCTGCTCGG 660
QY 661 CGCGTCTCGCTGCGCTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTGCGCTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTGCGATGCTGTTGCCAAGATCAGATGGCGCTGGGCGCAA 840
Db 781 ATGAGGGCATGTTCCCACTGCGATGCTGTTGCCAAGATCAGATGGCGCTGGGCGCAA 840
QY 841 TGGCGGCCATTAACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATAGC 900
Db 841 TGGCGGCCATTAACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATAGC 900
QY 901 AGGATACCGAAGACAGCTCATGTTATATCCCGCCGTTTAAACCAATCAAAACAGGATTTTC 960
Db 901 AGGATACCGAAGACAGCTCATGTTATATCCCGCCGTTTAAACCAATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACGAGCTGGACCGCTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGCTGGACCGCTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAAGAAAAACCACTTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAAGAAAAACCACTTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGCGAGCTGGCAGCAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGCGAGCTGGCAGCAGGTTT 1140
QY 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG 1200
Db 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG 1200
QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGCGGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGCGGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGCAGCCATTCGGAAGCTGTGGTATGGCTGTGAGGTCGTAATFCACTGATTAATTCG 1320
Db 1261 TCAGCAGCCATTCGGAAGCTGTGGTATGGCTGTGAGGTCGTAATFCACTGATTAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATATGTTTTTGGCGCGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATATGTTTTTGGCGCGACATCATACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTGCAATTAATCATCGGCTCGTATAATGTGTGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTGCAATTAATCATCGGCTCGTATAATGTGTGA 1440
QY 1441 ATTGTGAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTCAAGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTCAAGA 1500
QY 1501 GCACCTTCCAAACAGGACCATAGATTATGAAAACTGAAGAAAGTTAACTGGTAACTCTGG 1560

Db 1501 GCACCTTCCAAACAGGACCAATAGATTATGAAAAATCGAAGAGGTAAACTGGTAACTCGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGGAAGAGAAATCCCAACAGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGGAAGAGAAATCCCAACAGTT 1680
Qy 1681 GCGCAACTGCGGATGCGCTGACATATCTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGCGGATGCGCTGACATATCTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCGCGCTGTTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCGCGCTGTTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACCGCAAGCTGTATACCGGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACCGCAAGCTGTATACCGGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGTCGCAACCGCCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGTCGCAACCGCCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGTAAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTACCTGCGCGCTGATGCTGCTGACGCGGGTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTGCGCGCTGATGCTGCTGACGCGGGTATGCGTTCAAG 2040
Qy 2041 TATGAAACGCGCAAGTACGATTAAGACCTGCGCGCTGGATACGCTGGCGCAAGCG 2100
Db 2041 TATGAAACGCGCAAGTACGATTAAGACCTGCGCGCTGGATACGCTGGCGCAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGTTGACCTGATTAAACAAACACACATGAATGACAGACCGCATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAAACAAACACACATGAATGACAGACCGCATTAC 2160
Qy 2161 TCCATCGAAGAGCTGCTTTAATAAGGCGAAACAGCGATGACATCAACGCGCGCTGG 2220
Db 2161 TCCATCGAAGAGCTGCTTTAATAAGGCGAAACAGCGATGACATCAACGCGCGCTGG 2220
Qy 2221 GCATGCTCCAACTCGACACAGCAAGTGAATTTATGTTGTAACGCTACTGCCACCTTC 2280
Db 2221 GCATGCTCCAACTCGACACAGCAAGTGAATTTATGTTGTAACGCTACTGCCACCTTC 2280
Qy 2281 AAGGGTCAACCATCCAAACCGCTTGGTGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGCTTGGTGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGGCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGGCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2460
Qy 2461 TTGCGGAAAGATCCACGTATTGCGCCACCATGAAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGCGGAAAGATCCACGTATTGCGCCACCATGAAACCGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCGCGAGATGTCGCTTTCTGTTATCGCTGCGTATCGGCGATCAAGGCC 2580
Db 2521 CCGAACATCCGCGAGATGTCGCTTTCTGTTATCGCTGCGTATCGGCGATCAAGGCC 2580
Qy 2581 GCCAGCGGTGTCAGATGTCGATGAAGCCCTGAAAGACGCGGACAGCTAATTCAGGCTCG 2640
Db 2581 GCCAGCGGTGTCAGATGTCGATGAAGCCCTGAAAGACGCGGACAGCTAATTCAGGCTCG 2640

Qy 2641 AACAAACAACATATAACAATAACAACCTCGGGATCGAGGGAAGGATTTCAAGATTC 2700
Db 2641 AACAAACAACATATAACAATAACAACCTCGGGATCGAGGGAAGGATTTCAAGATTC 2700
Qy 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709
RESULT 7
ABL53238
ID ABL53238 standard; DNA; 8101 BP.
XX ABL53238;
AC ABL53238;
XX 17-JUN-2002 (first entry)
DT 17-JUN-2002 (first entry)
XX Nucleotide sequence of pMAL-Hs Cdc25C.
DE Cdc25C phosphatase; amine; Cdc25 phosphatase; spontaneous alopecia;
XX alopecia; proliferative disease; parasitic disease; viral infection;
KW neurodegenerative disease; myopathy; angiogenesis; psoriasis; restenosis;
KW vitamin K; ss.
XX Synthetic.
OS Homo sapiens.
XX Key Location/Qualifiers
FT CDS 1528..2677
FT FT /*tag= b
FT FT /note= "encodes maltose binding protein (MBP)"
FT CDS 2713..4134
FT FT /*tag= a
FT FT /note= "ORF of human Cdc25C"
XX FR2812198-A1.
XX 01-FEB-2002.
XX 28-JUL-2000; 2000FR-00009900.
XX 28-JUL-2000; 2000FR-00009900.
XX (SCRC) SCRAS SOC CONSEILS RECH & APPL SCI.
XX Prevost G, Brezak PMC, Galcera CMO, Thurieu C, Goubin GF;
PI Ducommun B;
XX WPI; 2002-166066/22.
XX cdc25 Phosphatase inhibiting medicaments, useful e.g. for treating tumor
PT diseases, viral infections, neurodegenerative disease or alopecia,
PT containing new or known aryl-substituted amines.
XX Disclosure; Page 18-22; 39pp; French.
XX The present sequence represents pMAL-HsCdc25C, a plasmid encoding human
CC Cdc25C phosphatase. The specification describes the use of aryl-
CC substituted secondary or tertiary amines for the production of
CC medicaments for inhibiting Cdc25 phosphatases, especially Cdc25C
CC phosphatases. The amines of the invention are used for treating
CC spontaneous alopecia or alopecia induced by exogenous products or
CC radiation. They are also used for treating tumoral or non-tumoral
CC proliferative diseases, parasitic diseases, viral infections,
CC neurodegenerative diseases or myopathy. Examples of non-tumoral
CC proliferative diseases are e.g. angiogenesis, psoriasis or restenosis.
CC The amines are also useful for inhibiting the proliferation of
CC microorganisms (especially yeasts) and in the treatment of all disorders
CC conventionally treated with vitamin K or its derivatives
XX Sequence 8101 BP; 2127 A; 1987 C; 2121 G; 1866 T; 0 U; 0 Other;

Query Match 82.0%; Score 2705.8; DB 6; Length 8101; Best Local Similarity 99.9%; Pred. No. 0; Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;									
Qy	1	CCGACACCATCGAATGGTGTGCAAAACCTTTTCGGGTATGGCATATGATAGCCCGGAAGAGA	60						
Db	1	CCGACACCATCGAATGGTGTGCAAAACCTTTTCGGGTATGGCATATGATAGCCCGGAAGAGA	60						
Qy	61	GTCAATTACAGGTGTGAATGTGAACCAAGTAACGTTATACGATGTTCGACAGTATGCCG	120						
Db	61	GTCAATTACAGGTGTGAATGTGAACCAAGTAACGTTATACGATGTTCGACAGTATGCCG	120						
Qy	121	GTGCTCTTATCAGACCGCTTTCCCGGTGGTGAACAGCCAGTATGATGCTGCGGAAA	180						
Db	121	GTGCTCTTATCAGACCGCTTTCCCGGTGGTGAACAGCCAGTATGATGCTGCGGAAA	180						
Qy	181	CGCGGAAAAGTGAAGCGCGATGCGGAGCTGAATTAACATCCCAACCGCTGGCAC	240						
Db	181	CGCGGAAAAGTGAAGCGCGATGCGGAGCTGAATTAACATCCCAACCGCTGGCAC	240						
Qy	241	AACAACTGGCGGCAAAACAGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300						
Db	241	AACAACTGGCGGCAAAACAGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300						
Qy	301	ACGCGCGTCCGAAATGTTCGGCGCATTAATCTCGCGCGATCAACTGGGTGCCAGC	360						
Db	301	ACGCGCGTCCGAAATGTTCGGCGCATTAATCTCGCGCGATCAACTGGGTGCCAGC	360						
Qy	361	TGTTGGTGTGATGTGATGAGCGGCGTGAAGCGCTGAAAGCGCGGTGCAACAATC	420						
Db	361	TGTTGGTGTGATGTGATGAGCGGCGTGAAGCGCTGAAAGCGCGGTGCAACAATC	420						
Qy	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGGTGATGACCAAGATGCCA	480						
Db	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGGTGATGACCAAGATGCCA	480						
Qy	481	TTGCTGTGGAAGCTGCCTGCATTAATGTTCGGCGCTTAATTTCTGATGCTCTGACAGA	540						
Db	481	TTGCTGTGGAAGCTGCCTGCATTAATGTTCGGCGCTTAATTTCTGATGCTCTGACCCA	540						
Qy	541	CACCCATCAACAGTATTAATTTCTCCATGAGACGGTACGGACTGGCGGTGGAGCATC	600						
Db	541	CACCCATCAACAGTATTAATTTCTCCATGAGACGGTACGGACTGGCGGTGGAGCATC	600						
Qy	601	TGTTGGTGTGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCGTCTCG	660						
Db	601	TGTTGGTGTGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCGTCTCG	660						
Qy	661	CGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
Db	661	CGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
Qy	721	CGGAACGGGAAGGACATGGAGTGCATGTCCGGTTTTCNACAAACCATGCAATGCTGA	780						
Db	721	CGGAACGGGAAGGACATGGAGTGCATGTCCGGTTTTCNACAAACCATGCAATGCTGA	780						
Qy	781	ATGAGGGCATGCTTCCCACTGCGATGCTGGTGGCCAAAGCATCAGATGGCGCTGGCGCAA	840						
Db	781	ATGAGGGCATGCTTCCCACTGCGATGCTGGTGGCCAAAGCATCAGATGGCGCTGGCGCAA	840						
Qy	841	TGCGCGCATTAACAGGTCCGGCTGGCGGTGGTGGGATATCTCGGTAGTGGGATAGC	900						
Db	841	TGCGCGCATTAACAGGTCCGGCTGGCGGTGGTGGGATATCTCGGTAGTGGGATAGC	900						
Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATCAACAGGATTTTC	960						
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATCAACAGGATTTTC	960						
Qy	961	GCCTGCTGGGCAACACAGCTGGAGCCGCTTCTCAGGCGCAGGCGGTGA	1020						
Db	961	GCCTGCTGGGCAACACAGCTGGAGCCGCTTCTCAGGCGCAGGCGGTGA	1020						
Qy	1021	AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA	1080						

Db	1021	AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA	1080						
Qy	1081	CGCAAAACCGCTCTCTCCCGCGCTTGGCCGATTCATTAATCAGCTGGCAGCAGCAGGTTT	1140						
Db	1081	CGCAAAACCGCTCTCTCCCGCGCTTGGCCGATTCATTAATCAGCTGGCAGCAGCAGGTTT	1140						
Qy	1141	CCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200						
Db	1141	CCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200						
Qy	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACGGTGACCAATGCTTCGCG	1260						
Db	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACGGTGACCAATGCTTCGCG	1260						
Qy	1261	TCAGGAGCCATCCGAAAGCTGTGCTGTGCTGCTGAGCTGCTGCTGCTGCTGCTGCTG	1320						
Db	1261	TCAGGAGCCATCCGAAAGCTGTGCTGTGCTGCTGAGCTGCTGCTGCTGCTGCTGCTG	1320						
Qy	1321	TGTCGCTCAAGCGCAGCTCCCGTTCTGGATTAATGTTTTTGGCGGAGCATCAATAACGTT	1380						
Db	1321	TGTCGCTCAAGCGCAGCTCCCGTTCTGGATTAATGTTTTTGGCGGAGCATCAATAACGTT	1380						
Qy	1381	CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGCGA	1440						
Db	1381	CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGCGA	1440						
Qy	1441	ATTGTGAGCGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTTTCAGCA	1500						
Db	1441	ATTGTGAGCGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTTTCAGCA	1500						
Qy	1501	GCATTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAAATCTGG	1560						
Db	1501	GCATTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAAATCTGG	1560						
Qy	1561	ATTAACCGCGATAAGGCTATAACGGTCTCGCTGAGTTCGCTGAGAAATTCGAGAAAGAT	1620						
Db	1561	ATTAACCGCGATAAGGCTATAACGGTCTCGCTGAGTTCGCTGAGAAATTCGAGAAAGAT	1620						
Qy	1621	ACCGAAATTAAGTCAACGGTTCGAGCATCCGATATACTGGAAGAGAAATTCGCCAGGTT	1680						
Db	1621	ACCGAAATTAAGTCAACGGTTCGAGCATCCGATATACTGGAAGAGAAATTCGCCAGGTT	1680						
Qy	1681	CGCGCACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740						
Db	1681	CGCGCACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740						
Qy	1741	GCTCAATCTGGCGTGTGGCTGAAATCACCCCGGACAAAGCGTTCAGGACAGCTGTAT	1800						
Db	1741	GCTCAATCTGGCGTGTGGCTGAAATCACCCCGGACAAAGCGTTCAGGACAGCTGTAT	1800						
Qy	1801	CCGTTTACCTGGGATGCGGTACGTTTACAAACGGCAAGCTGATTGCTTACCGATCGCTGT	1860						
Db	1801	CCGTTTACCTGGGATGCGGTACGTTTACAAACGGCAAGCTGATTGCTTACCGATCGCTGT	1860						
Qy	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA	1920						
Db	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA	1920						
Qy	1921	GAGATCCCGCGCTGGATTAAGAACTGAAAGGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980						
Db	1921	GAGATCCCGCGCTGGATTAAGAACTGAAAGGAAAGGTAAAGAGCGCGCTGATGTTCAAC	1980						
Qy	1981	CTGCAAGAAACGCTCTTCCACCTGGCGCTGATGCTGCTGAGGGGGTTATGCGTTCAAG	2040						
Db	1981	CTGCAAGAAACGCTCTTCCACCTGGCGCTGATGCTGCTGAGGGGGTTATGCGTTCAAG	2040						
Qy	2041	TATGAAAAACGCAAGTACGACATTAAGACAGCTGGGCGTGGATAAACGCTGGCGCGAAAGCG	2100						
Db	2041	TATGAAAAACGCAAGTACGACATTAAGACAGCTGGGCGTGGATAAACGCTGGCGCGAAAGCG	2100						
Qy	2101	GGTCTGACCTTCTCGGTTGACCTGATTTAAAAACAACATGAATGACAGACCCGATTAC	2160						

```

Db 2101 GGTCTGACCTTCTCGTGGTGGACCTGATTAATAAACAACACATGATGATGACACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAATAAGCGCAACAGCGATGACCATCAACGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAATAAGCGCAACAGCGATGACCATCAACGCGCCGCTGG 2220
QY 2221 GCATGGTCAACATCGACACACGCAAGAGTGAATATGTTGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACACGCAAGAGTGAATATGTTGTAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTCTGTCGCTGCTGAGGCGAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTCTGTCGCTGCTGAGGCGAGGTATTAACGCGCCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAGATCCACGCTATTGCGCCACCATGGAAACGCCGACAGAAAGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGCTATTGCGCCACCATGGAAACGCCGACAGAAAGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCGCTTTCTGTTGTTATGCGGTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCGCTTTCTGTTGTTATGCGGTGCTGCTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATCGAGTCTG 2640
Db 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATCGAGTCTG 2640
QY 2641 AACACAACAACAATAACAATAACAACAACCTCGGATCGAGGGAAGATTTTCAAGATTC 2700
Db 2641 AACACAACAACAATAACAATAACAACAACCTCGGATCGAGGGAAGATTTTCAAGATTC 2700
QY 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709

RESULT 8
AD023608
ID AD023608 standard; DNA; 7259 BP.
AC
XX
AC AD023608;
XX
DT 01-JUL-2004 (first entry)
XX
DE DNA encoding MBP-Toxop30del18(82-294aa) fusion protein.
XX
KW P30 antigen; Toxop30del13C; Toxop30del12C; Toxop30 MIX1;
KW MBP-Toxop30del13C(52-300aa); MBP-Toxop30del14C(52-294aa); MBP-Toxop30MIX1;
KW Toxoplasma gondii; ds; gene.
XX
OS Toxoplasma gondii.
OS Synthetic.
XX
PN US2004067239-A1.
XX
PD 08-APR-2004.
XX
PF 02-OCT-2002; 2002US-00263153.
XX
PR 02-OCT-2002; 2002US-00263153.
XX
PA (MAIN/) MAINE G T.
PA (PATE/) PATEL C B.
PA (GINS/) BLIESE S R.
PA (BLIE/) BLIESE T R.
XX

```

```

PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX WPI; 2004-304563/28.
DR P-PSDB; ADO23609.
XX
PT Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxop30del13C,
PT Toxop30del12C, Toxop30 MIX1, useful for detecting presence of IgM or IgG
PT antibodies to Toxoplasma gondii.
XX
XX Example 2; Fig 22; 114pp; English.
XX
CC The invention relates to a purified P30 antigen (I) chosen from 3 fully
CC defined Toxop30del13C, Toxop30del12C and Toxop30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxop30del12C P30 antigen sequence, or comprising an amino acid sequence in
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxop30del13C P30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxop30del13C(52-300aa), MBP-Toxop30del14C(52-294aa) and MBP-Toxop30MIX1
CC fusion proteins. (I) is useful for detecting the presence of IgM
CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC Toxop30 fusion protein of the invention.
XX
SQ Sequence 7259 BP; 1870 A; 1822 C; 1926 G; 1641 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7259;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGAAGAGA 60
Db 1 CCACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGAAGAGA 60
QY 61 GTCATTTACGGGTGGTGAATGTGAACCACTGATACGATGTGCGAGATATGCCG 120
Db 61 GTCATTTACGGGTGGTGAATGTGAACCACTGATACGATGTGCGAGATATGCCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCCGGTGTGTAACCAAGCCAGCCACCGTTTTCGCAAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCCGGTGTGTAACCAAGCCAGCCACCGTTTTCGCAAAA 180
QY 181 CGCGGGAAGAGTGAAGCGCGATGCGGAGCTGAATTACATTCGCAACCGCGTGGCAC 240
Db 181 CGCGGGAAGAGTGAAGCGCGATGCGGAGCTGAATTACATTCGCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGCAACACAGTCTGTTCTGATTTGGCGTTGCGACCTCGAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAACACAGTCTGTTCTGATTTGGCGTTGCGACCTCGAGTCTGGCCCTGC 300
QY 301 ACGCGCGTGCAGAAATGTCGGCGCATTAATCTCGCGCGCATCAATCTGGGTGGCCAGCG 360
Db 301 ACGCGCGTGCAGAAATGTCGGCGCATTAATCTCGCGCGCATCAATCTGGGTGGCCAGCG 360
QY 361 TGGTGTGTGATGTAGAACGAGCGGTGAAAGCGGCTGAAAGCGGCGGTGCACAATC 420
Db 361 TGGTGTGTGATGTAGAACGAGCGGTGAAAGCGGCTGAAAGCGGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACGCGTCACTGATTAATCTATCCGCTGATTAATCTGATGATGATGATGATG 480
Db 421 TTCTCGCGCAACGCGTCACTGATTAATCTATCCGCTGATTAATCTGATGATGATGATGATG 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTATTTCTTGATGTCTGTGACACAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTATTTCTTGATGTCTGTGACACAGA 540
QY 541 CACCCATCAACAGTATTTCTCCCATGAAGACGGTACGGAGCTGGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTCTCCCATGAAGACGGTACGGAGCTGGGCGGTGGAGCATC 600

```

QY 601 TGGTCGCAATTGGGTCACACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAAGTTCTGTCTCGG 660
Db 601 TGGTCGCAATTGGGTCACACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGGCTCGGCTGGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGGCTCGGCTGGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAAACGGGAAGGCGATGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAATGCTGA 780
Db 721 CGGAAACGGGAAGGCGATGGAGTGCCATGTCCGGTTTTCAACAAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTGGCATGCTGGTTCGCAACGATCAGATGCGCTGGCGCAA 840
Db 781 ATGAGGGCATGTTCCCACTGGCATGCTGGTTCGCAACGATCAGATGCGCTGGCGCAA 840
QY 841 TGGCGGCCATTACCGAGTCCGGGCTGCGGCTGGTGCGGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGGCCATTACCGAGTCCGGGCTGCGGCTGGTGCGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAAGACAGCTCATGTTATATCCGCGGTTAACCAACCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAAGACAGCTCATGTTATATCCGCGGTTAACCAACCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACAGGTGACCGCTTGCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACAGGTGACCGCTTGCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTCGCGCTCCTCACTGGTGAAAGAAACACACCTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTCGCGCTCCTCACTGGTGAAAGAAACACACCTGGCGGCCAATA 1080
QY 1081 CGCAACCGCCTCTCCCGCGGCTGGCGGATTCATTAATGCACTGGCAGACAGGTTT 1140
Db 1081 CGCAACCGCCTCTCCCGCGGCTGGCGGATTCATTAATGCACTGGCAGACAGGTTT 1140
QY 1141 CCCGACTGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTCTCATTTAG 1200
Db 1141 CCCGACTGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTCTCATTTAG 1200
QY 1201 GCACAATTCATGTTTCACAGCTTATCATGACCTGCAGGTGACCAATGCTTCGGG 1260
Db 1201 GCACAATTCATGTTTCACAGCTTATCATGACCTGCAGGTGACCAATGCTTCGGG 1260
QY 1261 TCAGGCGCCATCGGAAGCTGTGGTATGCTGTGAGGTGCTGATAATCACTGCATAATTCG 1320
Db 1261 TCAGGCGCCATCGGAAGCTGTGGTATGCTGTGAGGTGCTGATAATCACTGCATAATTCG 1320
QY 1321 TGTGCTCAAGCGGCATCTCCGTTCTGGATAATGTTTTTGGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGGCATCTCCGTTCTGGATAATGTTTTTGGCGCGACATCAATACGGTT 1380
QY 1381 CTGGCAATATCTCAAAATGAGCTGTCGACATTAATCATCGGCTCGTATATGTCGA 1440
Db 1381 CTGGCAATATCTCAAAATGAGCTGTCGACATTAATCATCGGCTCGTATATGTCGA 1440
QY 1441 ATTGTGAGCGGATAACAATTTTCACAGAAACAGCGAGTCCGTTTAGTGTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAATTTTCACAGAAACAGCGAGTCCGTTTAGTGTTCACGA 1500
QY 1501 GCATCTCAACAAAGGACCATAGCATATGAAAACTGAAAGAGGTAAACTGGTAATCTGG 1560
Db 1501 GCATCTCAACAAAGGACCATAGCATATGAAAACTGAAAGAGGTAAACTGGTAATCTGG 1560
QY 1561 ATTACGGGGATAAAGGCTATTAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTACGGGGATAAAGGCTATTAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGGCCCTGACATTAATCTTCTGGGCAACAGACCGCTTTGGTGGCTAC 1740

RESULT 9
ADO23613
ID ADO23613 standard; DNA; 7322 BP.
XX
AC ADO23613;

Db 1681 GCGCAACTGGCGATGGCCCTTGA CATTATCTTTGGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGAT 1800
Db 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTAAACGCGAAGCTGATTGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTAAACGCGAAGCTGATTGCTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGCGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGCGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTAAGACGTTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGTTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACAACATCAATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACAACATCAATGACAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAAAGTGCCTTTAATAAGCGAAACAGCGATGA CCAATCAAGCGCCGTGG 2220
Db 2161 TCCATCGCAGAAAGTGCCTTTAATAAGCGAAACAGCGATGA CCAATCAAGCGCCGTGG 2220
QY 2221 GCATGTCACACATCGACACAGCAAGTGAATTAATGCTGTAACGCTGACCGACCTTC 2280
Db 2221 GCATGTCACACATCGACACAGCAAGTGAATTAATGCTGTAACGCTGACCGACCTTC 2280
QY 2281 AAGGTCACACCATCAAAACCGTTGTTGGCGTGTGAGCGAGGTATTAAGCGCCCAAGT 2340
Db 2281 AAGGTCACACCATCAAAACCGTTGTTGGCGTGTGAGCGAGGTATTAAGCGCCCAAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGTTAATAAGACAAACCCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGTTAATAAGACAAACCCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGGAAAGATCCACGTATTTGCGCCACCATGGAAAAACGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGGAAAGATCCACGTATTTGCGCCACCATGGAAAAACGCCAGAAAGGTGAAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACCAACAACAATAACAATAAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACCAACAACAATAACAATAAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700

XX 01-JUL-2004 (first entry)
 XX DNA encoding MBP-ToxoP30del110(52-284aa) fusion protein.
 XX
 KW P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
 KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
 KW Toxoplasma gondii; da; gene.
 XX
 OS Toxoplasma gondii.
 OS Synthetic.
 XX US2004067239-A1.
 XX
 XX 08-APR-2004.
 XX
 XX 02-OCT-2002; 2002US-00263153.
 XX
 XX 02-OCT-2002; 2002US-00263153.
 XX (MAIN/) MAINE G T.
 XX (PATE/) PATEL C B.
 XX (GINS/) GINSBURG S R.
 XX (BLIE/) BLIESE T R.
 XX
 FI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
 XX
 XX WPI; 2004-304563/28.
 DR P-PSDB; ADO23614.
 XX
 XX Novel purified polypeptide having sequence identity to amino acid
 PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
 PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
 PT antibodies to Toxoplasma gondii.
 XX
 XX Example 2; Fig 25; 114pp; English.
 XX
 CC The invention relates to a purified P30 antigen (I) chosen from 3 fully
 CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
 CC amino acids added to the C-terminus of the amino acid sequence of
 CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
 CC which at least one of the five C-terminal cysteine amino acids of the
 CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
 CC with alanine, or comprising the amino acid sequence chosen from MBP-
 CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
 CC fusion proteins. (I) is useful for detecting the presence of IgM
 CC antibodies to Toxoplasma gondii in a test sample, which involves
 CC contacting the test sample suspected of containing the IgM antibodies
 CC with a composition comprising (I) and detecting the presence of (I)/IgM
 CC antibody complexes. The present sequence represents DNA encoding a MBP-
 CC ToxoP30 fusion protein of the invention.
 XX
 XX Sequence 7322 BP; 1887 A; 1844 C; 1934 G; 1657 T; 0 U; 0 Other;
 QQ
 Query Match 81.5%; Score 2690.4; DB 12; Length 7322;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CCGACACCATCGAATGTGCAAAACCTTTCGCGGTATGGATGATAGCGCCCGAAGAGA 60
 DB 1 CCGACACCATCGAATGTGCAAAACCTTTCGCGGTATGGATGATAGCGCCCGAAGAGA 60
 QY 61 GTCAATTCAGGGTGGTGAATGTGAACACAGTATACGATGTGCGAGAGTATGCCG 120
 DB 61 GTCAATTCAGGGTGGTGAATGTGAACACAGTATACGATGTGCGAGAGTATGCCG 120
 QY 121 GTGTCTTTATCAGACCGTTTCCCGCGTGTGTGAACAGGCCAGCCAGCTTTCGCGAAAA 180
 DB 121 GTGTCTTTATCAGACCGTTTTCGCGCGTGTGTGAACAGGCCAGCCAGCTTTCGCGAAAA 180
 QY 181 CCGGGGAAAAAGTGAAGCGCGCATGCGGAGCTGAATTACATTCCCAACCGGTTGGCAC 240
 DB 181 CCGGGGAAAAAGTGAAGCGCGCATGCGGAGCTGAATTACATTCCCAACCGGTTGGCAC 240

QY 241 AACAACTGGCGGCGCAACAGTTCGTTGCTGATTGSCGTTGCCACCTCCAGTCTGSCCTGC 300
 DB 241 AACAACTGGCGGCGCAACAGTTCGTTGCTGATTGSCGTTGCCACCTCCAGTCTGSCCTGC 300
 QY 301 ACSCGCGTCCCAAAATTGTTCGCGCGCATTAATCTCCGCGCGATCAACTGGGTGCCAGCG 360
 DB 301 ACSCGCGTCCCAAAATTGTTCGCGCGCATTAATCTCCGCGCGATCAACTGGGTGCCAGCG 360
 QY 361 TGGTGGTGTGATGGTAGAAGCGGCGTGAAGCCTGTAAAGCGGCGGTGCACAATC 420
 DB 361 TGGTGGTGTGATGGTAGAAGCGGCGTGAAGCCTGTAAAGCGGCGGTGCACAATC 420
 QY 421 TTCTTCGCGCAACGCGTCAGTGGGCTGATCAATTAACCTATCCGCTGATGACCAAGATGCCA 480
 DB 421 TTCTTCGCGCAACGCGTCAGTGGGCTGATCAATTAACCTATCCGCTGATGACCAAGATGCCA 480
 QY 481 TTGCTGTGGAAGTGCCTGCACATAAATGTTCCGGCGTTAATTTCTTGATGTCTCTGACCAGA 540
 DB 481 TTGCTGTGGAAGTGCCTGCACATAAATGTTCCGGCGTTAATTTCTTGATGTCTCTGACCAGA 540
 QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGATC 600
 DB 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGATC 600
 QY 601 TGGTTCGCATTGGGTCAACAGCAAAATCGCGCTGTTAGCGGCGCCATTAAAGTTCTGTCGG 660
 DB 601 TGGTTCGCATTGGGTCAACAGCAAAATCGCGCTGTTAGCGGCGCCATTAAAGTTCTGTCGG 660
 QY 661 CGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
 DB 661 CGCGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
 QY 721 CGGAAACGGGAAGCGGACTGAGTGCATGTCGCGTTTCAACAAACCATCAAAATGCTGA 780
 DB 721 CGGAAACGGGAAGCGGACTGAGTGCATGTCGCGTTTCAACAAACCATCAAAATGCTGA 780
 QY 781 ATGAGGCGCATCGTTCCCACTCGCATGCTGTTGCCAACAGTACAGTGGCGCTGGGCGCAA 840
 DB 781 ATGAGGCGCATCGTTCCCACTCGCATGCTGTTGCCAACAGTACAGTGGCGCTGGGCGCAA 840
 QY 841 TGGCGCCATTACCGAGTCCGGCTCGCGCTGCGGATATCTCGGTAGTGGGATACG 900
 DB 841 TGGCGCCATTACCGAGTCCGGCTCGCGCTGCGGATATCTCGGTAGTGGGATACG 900
 QY 901 ACCATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAACATCAAAACAGGATTTTC 960
 DB 901 ACCATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAACATCAAAACAGGATTTTC 960
 QY 961 GCCTGTGCGGCGCAAAACAGCGTGCAGCGCTGCTGCTCAACTCTCTCAGGGCGAGCGGTGA 1020
 DB 961 GCCTGTGCGGCGCAAAACAGCGTGCAGCGCTGCTGCTCAACTCTCTCAGGGCGAGCGGTGA 1020
 QY 1021 AGGGCAATCAGCTGTTGCGCGTCTCACTGGTGAAGAAAAACACCTCGGCGCCCAATA 1080
 DB 1021 AGGGCAATCAGCTGTTGCGCGTCTCACTGGTGAAGAAAAACACCTCGGCGCCCAATA 1080
 QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGCGAGCTGGCAGCAGAGTTT 1140
 DB 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGCGAGCTGGCAGCAGAGTTT 1140
 QY 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTTAGTTAGTCACTCATTAG 1200
 DB 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTTAGTTAGTCACTCATTAG 1200
 QY 1201 GCACAATTCATGTTTGAAGCTTATCATGCTGACCGGTGCACCAATGCTTCTGGCG 1260
 DB 1201 GCACAATTCATGTTTGAAGCTTATCATGCTGACCGGTGCACCAATGCTTCTGGCG 1260
 QY 1261 TCAGGCGAGCCATCGGAAGCTGTGATGGCTGTGCGAGTGTGTAATCACTGCAATATTCG 1320
 DB 1261 TCAGGCGAGCCATCGGAAGCTGTGATGGCTGTGCGAGTGTGTAATCACTGCAATATTCG 1320

QY 1321 TGTGCTCAAGCGCAGCTCCCGTTCTGGATATATGTTTTTGGCCGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGCAGCTCCCGTTCTGGATATATGTTTTTGGCCGACATCATACGGTT 1380
QY 1381 CTGGCAATATTTCTGAAATAGCTGTTCACAAATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAATATTTCTGAAATAGCTGTTCACAAATTAATCATCGGCTCGTATATGTTGGA 1440
QY 1441 ATTGTAGCGGATTAACAATTTTCAACAGAAACAGCCAGTCGGTTTGGTTTTCAGCA 1500
Db 1441 ATTGTAGCGGATTAACAATTTTCAACAGAAACAGCCAGTCGGTTTGGTTTTCAGCA 1500
QY 1501 GCACCTTCAACACAGGACCATAGATTATGAAACTGAAGAGGTAAACTGGTATCTGG 1560
Db 1501 GCACCTTCAACACAGGACCATAGATTATGAAACTGAAGAGGTAAACTGGTATCTGG 1560
QY 1561 ATTTAAGCGGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTTAAGCGGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATATAAATCGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATATAAATCGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGCAACTGGCGATGGCCCTGACATTATCTCTGGGCACACGCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTATCTCTGGGCACACGCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTCTGGCTGAAATCAACCGGACAAAGCGTTCCAGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTCTGGCTGAAATCAACCGGACAAAGCGTTCCAGACAAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCGGTACGTTTACAAACGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGATGCGGTACGTTTACAAACGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
QY 1861 GAAGCGTTTATCGCTGATTTATTAACAAGATCTGCTCCGACAAACCGCCCAAAACCTGGAA 1920
Db 1861 GAAGCGTTTATCGCTGATTTATTAACAAGATCTGCTCCGACAAACCGCCCAAAACCTGGAA 1920
QY 1921 GAGATCCCGCGCTCGATTAAGAACTGAAAGCGAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTCGATTAAGAACTGAAAGCGAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTTGCTGTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTTGCTGTGACGGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAACAAACACATGATGACACACCATTTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAACAAACACATGATGACACACCATTTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGCGCCGCTGG 2220
QY 2221 GCATGCTCAACATCGACACAGCAAGATTAATGTTGATGCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACAGCAAGATTAATGTTGATGCGGACCTTC 2280
QY 2281 AAGGTCACACATCAACCGTTCTGCGCTGAGCGGAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTCACACATCAACCGTTCTGCGCTGAGCGGAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTTCGAAAACATATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTTCGAAAACATATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCCGTAGCGCTGAAGTCTTTACGAGAAAG 2460

Db 2401 GAACGGTTAATAAGACAAACCGCTGGTCCGTAGCGCTGAAGCTTTACAGGAAAG 2460
QY 2461 TTGCGAAAGATCCACGTTATTGCCGCCACCATGGAACACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGCGAAAGATCCACGTTATTGCCGCCACCATGGAACACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACACTGTGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACACTGTGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG 2640
QY 2641 AACCAACAACAATAACAATAACAACAACCTCGGGATCGAGGGAGGATTTTCAAGATTC 2700
Db 2641 AACCAACAACAATAACAATAACAACAACCTCGGGATCGAGGGAGGATTTTCAAGATTC 2700
RESULT 10
ADO23603
ID ADO23603 standard; DNA; 7352 Bp.
XX
AC ADO23603;
XX
DT 01-JUL-2004 (first entry)
XX
DE DNA encoding MBP-ToxoP30del14C(52-294aa) fusion protein.
XX
KW P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; gene.
XX
OS Toxoplasma gondii.
OS Synthetic.
XX
PN US2004067239-A1.
XX
PD 08-APR-2004.
XX
PF 02-OCT-2002; 2002US-00263153.
XX
PR 02-OCT-2002; 2002US-00263153.
XX
PA (MAIN/) MAINE G T.
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX
PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX
DR WPI; 2004-304563/28.
DR P-PSDB; ADO23604.
XX
PT Novel purified polypeptide having sequence identity to amino acid sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C, Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG antibodies to Toxoplasma gondii.
XX
PS Example 2; Fig 19; 114pp; English.
XX
CC The invention relates to a purified P30 antigen (I) chosen from 3 fully defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6 amino acids added to the C-terminus of the amino acid sequence of Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in which at least one of the five C-terminal cysteine amino acids of the amino acid sequence of Toxo30del13C P30 antigen sequence is substituted with alanine, or comprising the amino acid sequence chosen from MBP-Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1 fusion proteins. (I) is useful for detecting the presence of IgM antibodies to Toxoplasma gondii in a test sample, which involves contacting the test sample suspected of containing the IgM antibodies

CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC Toxop30 fusion protein of the invention.
XX
SQ Sequence 7352 BP; 1896 A; 1852 C; 1943 G; 1661 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7352;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy	1	CCGACACCATCGAATGGTGCRAAACCTTTCGGGTATGGCATGATAGGCCCGCGAAGAGA	60
Db	1	CCGACACCATCGAATGGTGCRAAACCTTTCGGGTATGGCATGATAGGCCCGCGAAGAGA	60
Qy	61	GTCAATTGAGGGTGGTGAATGTGAACCCAGTAAAGCTTATACGATGTGCGCAGATATGCCG	120
Db	61	GTCAATTGAGGGTGGTGAATGTGAACCCAGTAAAGCTTATACGATGTGCGCAGATATGCCG	120
Qy	121	GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACCCAGGCGAGCCACGTTTCTCGAATAA	180
Db	121	GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACCCAGGCGAGCCACGTTTCTCGAATAA	180
Qy	181	CGCGGGAAGTGAAGCGCGATGCGGAGCTGAATTAATCTCGCGCGATCAATCTCGGCAAA	240
Db	181	CGCGGGAAGTGAAGCGCGATGCGGAGCTGAATTAATCTCGCGCGATCAATCTCGGCAAA	240
Qy	241	AACAACTGGCGGCAAAACAGTCGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGCGCCCTGC	300
Db	241	AACAACTGGCGGCAAAACAGTCGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGCGCCCTGC	300
Qy	301	ACGCGCGTGCCTAATTTGTCGCGCGATTAATCTCGCGCGATCAATCTCGGTCGCGAGC	360
Db	301	ACGCGCGTGCCTAATTTGTCGCGCGATTAATCTCGCGCGATCAATCTCGGTCGCGAGC	360
Qy	361	TGTTGTTGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT	420
Db	361	TGTTGTTGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT	420
Qy	421	TTCTCGCGCAACCGTTCAGTGGGCTGATCAATTAATCTCGCGCGATCAATCTCGGTCGCGAGC	480
Db	421	TTCTCGCGCAACCGTTCAGTGGGCTGATCAATTAATCTCGCGCGATCAATCTCGGTCGCGAGC	480
Qy	481	TTGCTGTGGAAGCTGCTGCACTAATGTCGCGGTTAATTTCTGATGTCGATGTCGATGTCGAT	540
Db	481	TTGCTGTGGAAGCTGCTGCACTAATGTCGCGGTTAATTTCTGATGTCGATGTCGATGTCGAT	540
Qy	541	CACCCATCAACAGTATTTTCTCCCATGAAGACGTCGCGACTGGCGGTGGAGCATC	600
Db	541	CACCCATCAACAGTATTTTCTCCCATGAAGACGTCGCGACTGGCGGTGGAGCATC	600
Qy	601	TGTTGCGATTTGGTTCACAGCAATCGCGTGTAGCGGGCCCAATTAAGTTCTGCTCGG	660
Db	601	TGTTGCGATTTGGTTCACAGCAATCGCGTGTAGCGGGCCCAATTAAGTTCTGCTCGG	660
Qy	661	CGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	720
Db	661	CGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	720
Qy	721	CGGAAACGGGAGGCGACTGGAGTGCCTATGTCGCGTTTTCACCAACCATCAATGCTGA	780
Db	721	CGGAAACGGGAGGCGACTGGAGTGCCTATGTCGCGTTTTCACCAACCATCAATGCTGA	780
Qy	781	ATGAGGCGCATGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGGCGCTGGCGCAA	840
Db	781	ATGAGGCGCATGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGGCGCTGGCGCAA	840
Qy	841	TGCGCGCATTCAGGATCGCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGATACG	900
Db	841	TGCGCGCATTCAGGATCGCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG	900
Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAATCAACAGGATTTTC	960
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAATCAACAGGATTTTC	960

Qy	961	GCCTGCTGGGCGCAAAACAGCGTGCAGCGCTTGTGCTCAACTCTCTCAGGCGCAGCGGTGA	1020
Db	961	GCCTGCTGGGCGCAAAACAGCGTGCAGCGCTTGTGCTCAACTCTCTCAGGCGCAGCGGTGA	1020
Qy	1021	AGGCAATCAGCTGTTGCCCGTCTCACTGCTGAAAAGAAAACACACCTGCGCGCCCAATA	1080
Db	1021	AGGCAATCAGCTGTTGCCCGTCTCACTGCTGAAAAGAAAACACACCTGCGCGCCCAATA	1080
Qy	1081	CGMAACCGGCTCTCCCGCGGCTTGGCCGATTCATTAATGTCAGCTGGCAGCAGAGTTT	1140
Db	1081	CGMAACCGGCTCTCCCGCGGCTTGGCCGATTCATTAATGTCAGCTGGCAGCAGAGTTT	1140
Qy	1141	CCGACCTGGAAGCGGCGAGTGCAGCAACGCAATTAATGTCAGCTGGCAGCAGAGTTT	1200
Db	1141	CCGACCTGGAAGCGGCGAGTGCAGCAACGCAATTAATGTCAGCTGGCAGCAGAGTTT	1200
Qy	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCACGCTGCACCAATGCTTCTCGCG	1260
Db	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCACGCTGCACCAATGCTTCTCGCG	1260
Qy	1261	TCAGGACCGCATCGGAAGCTGTGATGGCTGTGAGGTGCTAAATCACTGCATAATTCG	1320
Db	1261	TCAGGACCGCATCGGAAGCTGTGATGGCTGTGAGGTGCTAAATCACTGCATAATTCG	1320
Qy	1321	TGTCGCTCAAGGCGCACTCCCGTCTGGATAATGTTTTTTCGCGCGACATCAATACGGTT	1380
Db	1321	TGTCGCTCAAGGCGCACTCCCGTCTGGATAATGTTTTTTCGCGCGACATCAATACGGTT	1380
Qy	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTAAATGTTGGGA	1440
Db	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTAAATGTTGGGA	1440
Qy	1441	ATTGTGAGCGGATTAACAAATTTTACACAGGAAACAGCCAGTCCGTTTGGTGGTTCACGA	1500
Db	1441	ATTGTGAGCGGATTAACAAATTTTACACAGGAAACAGCCAGTCCGTTTGGTGGTTCACGA	1500
Qy	1501	GCACTTTCACCAACAGGACCATAGATTATGAAATCTGAAAGGTAACTGGTAACTCTGG	1560
Db	1501	GCACTTTCACCAACAGGACCATAGATTATGAAATCTGAAAGGTAACTGGTAACTCTGG	1560
Qy	1561	ATTAAACCGGATTAAGGCTATACCGTCTCGTGAAGTTCGTAAGAAATTCGAGAAGAT	1620
Db	1561	ATTAAACCGGATTAAGGCTATACCGTCTCGTGAAGTTCGTAAGAAATTCGAGAAGAT	1620
Qy	1621	ACCGGAATTAAGTTCACCGTTCGGAATCCGGATTAACCTGGAAGAGAAATTCACAGGTT	1680
Db	1621	ACCGGAATTAAGTTCACCGTTCGGAATCCGGATTAACCTGGAAGAGAAATTCACAGGTT	1680
Qy	1681	CGGCAACTGGCGGATGGCCCTGACATTTCTTCTGGGCAACACCGCTTGGTGGCTAC	1740
Db	1681	CGGCAACTGGCGGATGGCCCTGACATTTCTTCTGGGCAACACCGCTTGGTGGCTAC	1740
Qy	1741	GCTCAATCTGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1800
Db	1741	GCTCAATCTGGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1800
Qy	1801	CGGTTTACCTGGGATCGGTCAGTTCACAGGCAAGCTGATTCCTTACCGGATCGCTGTT	1860
Db	1801	CGGTTTACCTGGGATCGGTCAGTTCACAGGCAAGCTGATTCCTTACCGGATCGCTGTT	1860
Qy	1861	GAGCGCTTATCGCTGATTTTATACAAAGATCTGCTGCCGAACCCGCGCAAAACCTGGGAA	1920
Db	1861	GAGCGCTTATCGCTGATTTTATACAAAGATCTGCTGCCGAACCCGCGCAAAACCTGGGAA	1920
Qy	1921	GAGATCCCGCGCTGATTAAGAACTGAAAGGTAAGAGCGCGCTGATGTTTCAAC	1980
Db	1921	GAGATCCCGCGCTGATTAAGAACTGAAAGGTAAGAGCGCGCTGATGTTTCAAC	1980
Qy	1981	CTGCAAGAACCGTACCTTTCACCTGGCGGCTGATTCCTGCTGACGCGGGTTATGCTTCAAG	2040
Db	1981	CTGCAAGAACCGTACCTTTCACCTGGCGGCTGATTCCTGCTGACGCGGGTTATGCTTCAAG	2040

Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCACTGGCGGTGGAGCATC 600
Qy 601 TGGTCCGATTGGGTCAACAGCAAAATCGCGCTGTATAGCGGGCCCAATTAAGTTCTGCTCGG 660
Db 601 TGGTCCGATTGGGTCAACAGCAAAATCGCGCTGTATAGCGGGCCCAATTAAGTTCTGCTCGG 660
Qy 661 CGCGTCTCGCTGCGCTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTGCGCTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGACATCGAGTGCATGTCCGTTTCAACAAACCAATGCAATGCTGA 780
Db 721 CGGAACGGGAAGGACATCGAGTGCATGTCCGTTTCAACAAACCAATGCAATGCTGA 780
Qy 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTCGCAACGATCAGATGCGCTGGCGGCAA 840
Db 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTCGCAACGATCAGATGCGCTGGCGGCAA 840
Qy 841 TGGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATAG 900
Db 841 TGGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATAG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAACCACCATCAACACGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAACCACCATCAACACGATTTTC 960
Qy 961 GCCTGTGGGGCAAAACCAAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGTGGGGCAAAACCAAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTGCGGCTCTCACTGGTGAAGAAACCAACCTGCGGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGCGGCTCTCACTGGTGAAGAAACCAACCTGCGGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCTGGCAGCAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCTGGCAGCAGGTTT 1140
Qy 1141 CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTCATTAG 1200
Db 1141 CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTCATTAG 1200
Qy 1201 GCACAAATCTCATGTTGACAGCTTATCATGCACTGACGGTGACCAATGCTTCTGGG 1260
Db 1201 GCACAAATCTCATGTTGACAGCTTATCATGCACTGACGGTGACCAATGCTTCTGGG 1260
Qy 1261 TCAGGACGCCATCGAAAGCTGTGTTATGGCTGTGCGAGTGTAAATCACTGCAATAATTCG 1320
Db 1261 TCAGGACGCCATCGAAAGCTGTGTTATGGCTGTGCGAGTGTAAATCACTGCAATAATTCG 1320
Qy 1321 TGTGCTCAAGGGCACTCCCGTTCTGGATAATGTTTTTTCGCGGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGGGCACTCCCGTTCTGGATAATGTTTTTTCGCGGACATCAATACGGTT 1380
Qy 1381 CTGGCAATATCTGAATGAGCTGTGACAAATTAATCATCGGCTGTAATGTTGGA 1440
Db 1381 CTGGCAATATCTGAATGAGCTGTGACAAATTAATCATCGGCTGTAATGTTGGA 1440
Qy 1441 ATTGTAGCGGATAACAAATTTACACAGGAAACAGCCAGTCCGTTTAAAGTGTTCACGA 1500
Db 1441 ATTGTAGCGGATAACAAATTTACACAGGAAACAGCCAGTCCGTTTAAAGTGTTCACGA 1500
Qy 1501 GCATTTACCAACAGGACCATAGATTAATGAAACTGAAAGGTAATTCGTAATCTGG 1560
Db 1501 GCATTTACCAACAGGACCATAGATTAATGAAACTGAAAGGTAATTCGTAATCTGG 1560
Qy 1561 ATTAACGGGATTAAGGCTTAACGGTCTCGCTGAAGTCCGGTAAAGATTCGAGAAGAT 1620
Db 1561 ATTAACGGGATTAAGGCTTAACGGTCTCGCTGAAGTCCGGTAAAGATTCGAGAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACACAGGTT 1680

RESULT 12

AD023639

ID ADO23639 standard; DNA; 7370 BP.

Qy 1681 GCGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACAGACCGCTTTGGTGGGTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACAGACCGCTTTGGTGGGTAC 1740
Qy 1741 GCTCAATCTGGCCTGTGGCTGAAATCACCCCGGACAAAGGTTTCCAGGACAACTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTGGCTGAAATCACCCCGGACAAAGGTTTCCAGGACAACTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACCGGCAAGCTGATTTGTTTACCGGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACCGGCAAGCTGATTTGTTTACCGGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATTAACAAAGATCTGCTGCCGAAACCCGCAAAACCTCGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATTAACAAAGATCTGCTGCCGAAACCCGCAAAACCTCGGAA 1920
Qy 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTATCTTCACTGCGCGCTGATTTGCTGCTGACCGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTATCTTCACTGCGCGCTGATTTGCTGCTGACCGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGGTTGACCTGATTTAAACCAACACATGAATGCAGACACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTTAAACCAACACATGAATGCAGACACCGGATTAC 2160
Qy 2161 TCCATGCGAAGCTGCTTTTAAAGGCGAAACAGCGATGACCATTAACGCGCGCGTGG 2220
Db 2161 TCCATGCGAAGCTGCTTTTAAAGGCGAAACAGCGATGACCATTAACGCGCGCGTGG 2220
Qy 2221 GCATGGTCCAAACATCGACACCGCAAGTGAATTTATGGTGTAAACGGTACTGCGACGCTTC 2280
Db 2221 GCATGGTCCAAACATCGACACCGCAAGTGAATTTATGGTGTAAACGGTACTGCGACGCTTC 2280
Qy 2281 AAGGGTCAACCATCCAAACCGTTCTGCTGAGCGCAGGTATTAACGCGCGCGAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTCTGCTGAGCGCAGGTATTAACGCGCGCGAGT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAATACTATCTGCTGACTGATGAAGGTCG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAATACTATCTGCTGACTGATGAAGGTCG 2400
Qy 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGGCGAAAGATCCAGTATTCGCGCACATGGAACAAACCGCGAAGGTTGAAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTCGCGCACATGGAACAAACCGCGAAGGTTGAAATCATG 2520
Qy 2521 CCGAACATCCGCGAGATGTCGGTTTCTGGTATCCGCTGGTACTCGCGTGTATCAAGCC 2580
Db 2521 CCGAACATCCGCGAGATGTCGGTTTCTGGTATCCGCTGGTACTCGCGTGTATCAAGCC 2580
Qy 2581 GCCAGCGGTCGTGAGACTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTCGTGAGACTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACCAACAAACAAATTAACAACTCGGATCGAGGAAAGGATTTTCAGAAATTC 2700
Db 2641 AACCAACAAACAAATTAACAACTCGGATCGAGGAAAGGATTTTCAGAAATTC 2700

XX AC ADO23639;
XX DT 01-JUL-2004 (first entry)
XX DE DNA encoding MBP-ToxoP30MIX1 fusion protein.
XX KW P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; gene.
XX OS Toxoplasma gondii.
OS Synthetic.
XX PN US2004067239-A1.
XX PD 08-APR-2004.
XX PF 02-OCT-2002; 2002US-00263153.
XX PR 02-OCT-2002; 2002US-00263153.
XX PA (MAIN/) MAINE G T.
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX WPI; 2004-304563/28.
XX P-PSDB; ADO23640.
XX PT Novel purified polypeptide having sequence identity to amino acid
PT sequence of p30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of Igm or Igg
PT antibodies to Toxoplasma gondii.
XX Example 5; Fig 31; 114pp; English.
XX The invention relates to a purified p30 antigen (I) chosen from 3 fully
CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxo30del12C p30 antigen sequence, or comprising an amino acid sequence in
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxo30del13C p30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
CC fusion proteins. (I) is useful for detecting the presence of Igm
CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the Igm antibodies
CC with a composition comprising (I) and detecting the presence of (I)/Igm
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC ToxoP30 fusion protein of the invention.
XX SQ Sequence 7370 BP; 1900 A; 1856 C; 1950 G; 1664 T; 0 U; 0 Other;
Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
Best Local Similarity 99.8%; Pred No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
Qy 61 GTCAAATTCAGGTGGTGAATGTGAACCAACAGTATATACGATGTCGAGAGTATGCGG 120
Db 61 GTCAAATTCAGGTGGTGAATGTGAACCAACAGTATATACGATGTCGAGAGTATGCGG 120
Qy 121 GTGTCTCTTATCAGACCGGTTTCCCGCGTGGTGAACCAACAGCCAGCCAGCTTTCTCGGAAAA 180
Db 121 GTGTCTCTTATCAGACCGGTTTCCCGCGTGGTGAACCAACAGCCAGCCAGCTTTCTCGGAAAA 180
Qy 181 CGCGGAAAAAGTGAAGCGCGCATGGCGGAGCTGAATTTACATTTCCCAACCGCTGGCAC 240

Db 181 CGCGGAAAAAGTGAAGCGCGCATGGCGGAGCTGAATTTACATTTCCCAACCGCTGGCAC 240
Qy 241 AACAACTGGCGGGCAAAACAGTCTGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGCAAAACAGTCTGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGCGCGTTCGCAAAATTTGTCGCGCGGATTAATCTCTCGCGCGGATCAACTCGGGTGGCAGCG 360
Db 301 ACGCGCGTTCGCAAAATTTGTCGCGCGGATTAATCTCTCGCGCGGATCAACTCGGGTGGCAGCG 360
Qy 361 TGGTGGTGTGATGTAAGAAAGCGCGCTGGAAGCCTGTAAAGCGCGGTGCAACAATC 420
Db 361 TGGTGGTGTGATGTAAGAAAGCGCGCTGGAAGCCTGTAAAGCGCGGTGCAACAATC 420
Qy 421 TTCTCGCGCAACCGGCTCAGTGGGCTGATCACTAATCTATCCGCTGGATCACCAGGATGCCA 480
Db 421 TTCTCGCGCAACCGGCTCAGTGGGCTGATCACTAATCTATCCGCTGGATCACCAGGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGGCTTATTTCTTGATGTCTCTGACGAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGGCTTATTTCTTGATGTCTCTGACGAGA 540
Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Qy 601 TGGTGGCATTTGGGTCACCGCAAAATCGGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG 660
Db 601 TGGTGGCATTTGGGTCACCGCAAAATCGGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG 660
Qy 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTCCTCAACAACCATGCAATGCTGA 780
Db 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTCCTCAACAACCATGCAATGCTGA 780
Qy 781 ATGAGGGCATGCTTCCCACTCGCATGCTGTTGCTCAACGATCAGATGCGCTGGCGGCA 840
Db 781 ATGAGGGCATGCTTCCCACTCGCATGCTGTTGCTCAACGATCAGATGCGCTGGCGGCA 840
Qy 841 TGC CGCGCATTTACCGAGTCCCGGCTGCGGATATCTCGGTAGTGGGATAG 900
Db 841 TGC CGCGCATTTACCGAGTCCCGGCTGCGGATATCTCGGTAGTGGGATAG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACCAACCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACCAACCATCAACAGGATTTTC 960
Qy 961 GCCTGCTGGGGCAACCAACGAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGCTGGGGCAACCAACGAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTTGCCCGCTCTCACTGTTGAAAAGAAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCCGCTCTCACTGTTGAAAAGAAAACCAACCTGGCGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGTGGCAACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGTGGCAACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACCAATTAATGTAGTTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACCAATTAATGTAGTTAGTCACTCATTTAG 1200
Qy 1201 GCACAAATCTCATGTTTTCAGACGCTTATCATCGACTGACGGTGCACCAATGTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTTCAGACGCTTATCATCGACTGACGGTGCACCAATGTTCTGGCG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGGTATGCGCTGTGAGGTGTAATCACTGCAATTTTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGCGCTGTGAGGTGTAATCACTGCAATTTTCG 1320

Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTAATCACTGCATAATTTCG 1320
Qy 1321 TGTCTCTCAAGCGCACCTCCGCTTCTGGATATGTTTTTTTGGCCGACATCATACGGTT 1380
Db 1321 TGTCTCTCAAGCGCACCTCCGCTTCTGGATATGTTTTTTTGGCCGACATCATACGGTT 1380
Qy 1381 CTGGCAAAATATTCTGAAATAGCTCTGTGACAAATTAAATCATCGGCTCGTATAATGTGTGGA 1440
Db 1381 CTGGCAAAATATTCTGAAATAGCTCTGTGACAAATTAAATCATCGGCTCGTATAATGTGTGGA 1440
Qy 1441 ATTGTGAGCGGATACAAATTTTACACAGAGAAACAGCCAGTCGCTTTAGGTGTTTTCAGGA 1500
Db 1441 ATTGTGAGCGGATACAAATTTTACACAGAGAAACAGCCAGTCGCTTTAGGTGTTTTCAGGA 1500
Qy 1501 GCACCTTCAACCAAGGACCATAGATTATGAAACTGGAAGAGGTAAACTGCTAATCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGCATATGAAATCGAAGAGGTAAACTGCTAATCTGG 1560
Qy 1561 ATTAAACGGGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAAACGGGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTACACGTTGAGCATCCGATCCGATAAACTGGAAGAGAAATTCGAGAAAGAT 1680
Db 1621 ACCGGAATTAAGTACACGTTGAGCATCCGATAAACTGGAAGAGAAATTCGAGAAAGAT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGTGGCTGAAATCAATCAACCCGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGTGGCTGAAATCAATCAACCCGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACCTCGGATGCGTACGTTACAAACGCAAGCTGATGCTTACCGCATCGCTGTT 1860
Db 1801 CCGTTTACCTCGGATGCGTACGTTACAAACGCAAGCTGATGCTTACCGCATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCAAAACCTCGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCAAAACCTCGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAGAACTGAAGCGAAAGTAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAGAACTGAAGCGAAAGTAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGCGCTGATGCTGCTGACGGGGTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCTGATGCTGCTGACGGGGTATGCGTTCAAG 2040
Qy 2041 TATGAAAAACGCAAGTACGACATTAAGACGTTGGCGCTGGATACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGACGTTGGCGCTGGATACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCCTGGTTGACCTGATTAAAAACAAACACATGATGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCCTGGTTGACCTGATTAAAAACAAACACATGATGAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTTAAATAAGGCGAAACACGCGATGACCATCAACGGCCGTTG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAATAAGGCGAAACACGCGATGACCATCAACGGCCGTTG 2220
Qy 2221 GCATGTCCAAATCGACACAGCAAAAGTGAATTTATGTTGTAACGGTACTGCCGACCTTC 2280
Db 2221 GCATGTCCAAATCGACACAGCAAAAGTGAATTTATGTTGTAACGGTACTGCCGACCTTC 2280
Qy 2281 AAGGGTCAACATCAACACGCTTGGTGGCTGCTGAGCCGAGGTATTACGGCGCAGT 2340
Db 2281 AAGGGTCAACATCAACACGCTTGGTGGCTGCTGAGCCGAGGTATTACGGCGCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400

Qy 2401 GAACGGTGTATATAAGACAAACCCCTGGTGCCTAGCGTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAACGGTGTATATAAGACAAACCCCTGGTGCCTAGCGTGAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGGCGAAAGATCCACGCTATTGGCCGACCATGGAACCCGCAAGAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGCTATTGGCCGACCATGGAACCCGCAAGAGGTGAATCATG 2520
Qy 2521 CCGAACATCCGCGAGATGTCCGCTTTCTGGTATGCCGTGGGTACTCGCGTGAATCAAGGCC 2580
Db 2521 CCGAACATCCGCGAGATGTCCGCTTTCTGGTATGCCGTGGGTACTCGCGTGAATCAAGGCC 2580
Qy 2581 GCCAGCGGTGCTGACAGCTGTCGATGAAGCCCTGAAAGACCGCGACGATTAATTCGAGTCTG 2640
Db 2581 GCCAGCGGTGCTGACAGCTGTCGATGAAGCCCTGAAAGACCGCGACGATTAATTCGAGTCTG 2640
Qy 2641 AACAAACAAACAAATAACAATAACAACAACTCGGGATCGAGGGAAGGATTTTCAAAATTC 2700
Db 2641 AACAAACAAACAAATAACAATAACAACAACTCGGGATCGAGGGAAGGATTTTCAAAATTC 2700

RESULT 13
ADO23649
ID ADO23649 standard; DNA; 7370 BP.
XX ADO23649;
AC ADO23649;
XX ADO23649;
DT 01-JUL-2004 (first entry)
XX DNA encoding MBP-Toxop30MIX5 fusion protein.
DE P30 antigen; Toxo30del3C; Toxo30del12C; ToxoP30 MIX1;
KW MBP-Toxo30del3C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; gene.
XX Toxoplasma gondii.
OS Synthetic.
XX US2004067239-A1.
XX 08-APR-2004.
XX 02-OCT-2002; 2002US-00263153.
XX 02-OCT-2002; 2002US-00263153.
XX (MAIN/) MAINE G T.
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX WPI; 2004-304563/28.
DR P-PSDB; ADO23650.
XX Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del3C,
PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
PT antibodies to Toxoplasma gondii.
XX Example 5; Fig 35; 114pp; English.
PS The invention relates to a purified P30 antigen (I) chosen from 3 fully
XX defined Toxo30del3C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxo30del3C P30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxo30del3C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
CC fusion proteins. (i) is useful for detecting the presence of IgM

CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (i) and detecting the presence of (i)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC Toxop30 fusion protein of the invention.
XX
SQ Sequence 7370 BP; 1900 A; 1857 C; 1950 G; 1663 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCGACACCATCGATCGTGAACACCTTTCGGGTATGCGCATGATAGCGCCGGAGAGA 60
DB 1 CCGACACCATCGAATGGTGAACACCTTTCGGGTATGCGCATGATAGCGCCGGAGAGA 60
QY 61 GTCAAATTCAGGGTGGTGAATGTGAACACAGTAAACGTTATACGATGTCGACAGTATGCCG 120
DB 61 GTCAAATTCAGGGTGGTGAATGTGAACACAGTAAACGTTATACGATGTCGACAGTATGCCG 120
QY 121 GTGTCCTTATACAGACCGTTTCCCGCGTGGTGAACACAGGCCAGCCAGCTTTCGCGAAAA 180
DB 121 GTGTCCTTATACAGACCGTTTCCCGCGTGGTGAACACAGGCCAGCCAGCTTTCGCGAAAA 180
QY 191 CGCGGAAAAAGTGAAGCGCGGATGCGGAGCTGAATTAACATCCCAACCGCGTGGCAC 240
DB 191 CGCGGAAAAAGTGAAGCGCGGATGCGGAGCTGAATTAACATCCCAACCGCGTGGCAC 240
QY 241 AACAACTCGCGGGCAACAGTGTGCTGATTGGGGTTGCCACCTTCAGTCTGGGCCCTGC 300
DB 241 AACAACTCGCGGGCAACAGTGTGCTGATTGGGGTTGCCACCTTCAGTCTGGGCCCTGC 300
QY 301 ACGCCCGTTCGCAAAATGTGCGGGCGATTAAATCTCGCGCGATCAATCGGTCGCGACG 360
DB 301 ACGCCCGTTCGCAAAATGTGCGGGCGATTAAATCTCGCGCGATCAATCGGTCGCGACG 360
QY 361 TGGTGGTTCGATGTAGACAAGCGGGCTCGAAGCGCTGAAAGCGGGTGCACAAATC 420
DB 361 TGGTGGTTCGATGTAGACAAGCGGGCTCGAAGCGCTGAAAGCGGGTGCACAAATC 420
QY 421 TTCTCGCGCAACGCGTCACTAGTGGGCTGATCAATTAATCTCGCGCGATCAATCGGTCGCGACG 480
DB 421 TTCTCGCGCAACGCGTCACTAGTGGGCTGATCAATTAATCTCGCGCGATCAATCGGTCGCGACG 480
QY 481 TTGCTGTGGAAGCTGCCTGCAATAATGTTCGGCGTTATTTCTGATGTTCTCTGACCCAGA 540
DB 481 TTGCTGTGGAAGCTGCCTGCAATAATGTTCGGCGTTATTTCTGATGTTCTCTGACCCAGA 540
QY 541 CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGGCATGCGCGTGGAGCATC 600
DB 541 CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGGCATGCGCGTGGAGCATC 600
QY 601 TGGTCGCATGGGTCCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
DB 601 TGGTCGCATGGGTCCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGGTCTGGCTGGCTGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTCGGTCTGGCTGGCTGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAAACGGGAAGCGGACGTGAGTGCATGTCGGTTTTCACAAACCATGCAAAATGCTGA 780
DB 721 CGGAAACGGGAAGCGGACGTGAGTGCATGTCGGTTTTCACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTCCCACTGCGATCGTGGTTGCCAAGCATCAGATGCGCGTGGCGGCA 840
DB 781 ATGAGGGCATCGTCCCACTGCGATCGTGGTTGCCAAGCATCAGATGCGCGTGGCGGCA 840
QY 841 TCGCGGCCATTAACAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
DB 841 TCGCGGCCATTAACAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCCATCAAAAGGATTTTC 960

DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCCATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGGCAAAACAGCGCTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGGCGTGA 1020
DB 961 GCCTGCTGGGGCAAAACAGCGCTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGGCGTGA 1020
QY 1021 AGGGCAATCAGCTGTGTCGCCGTCTCACTGGTGAAGAAAAACCAACCTGCGGCCCAATA 1080
DB 1021 AGGGCAATCAGCTGTGTCGCCGTCTCACTGGTGAAGAAAAACCAACCTGCGGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATCATTAATGCAGCTGCGACAGAGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATCATTAATGCAGCTGCGACAGAGTTT 1140
QY 1141 CCCGACTTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
DB 1141 CCCGACTTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGCTGCAACCAATGCTTCGGG 1260
DB 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGCTGCAACCAATGCTTCGGG 1260
QY 1261 TCAGCAGCCATCGAAGCTGTGCTGATGCGTGTGAGCTCGTAATCACTGCAATAATTCG 1320
DB 1261 TCAGCAGCCATCGAAGCTGTGCTGATGCGTGTGAGCTCGTAATCACTGCAATAATTCG 1320
QY 1321 TGTGCTCAAGCGCACCTCCCGTCTTGGATAATGTTTTTGGCGCGACATCAATACGGTT 1380
DB 1321 TGTGCTCAAGCGCACCTCCCGTCTTGGATAATGTTTTTGGCGCGACATCAATACGGTT 1380
QY 1381 CTGGCAAAATCTTGAATAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGTGGA 1440
DB 1381 CTGGCAAAATCTTGAATAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGTGGA 1440
QY 1441 ATTGTGAGCGGATTAACAAATTTTCAACAGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
DB 1441 ATTGTGAGCGGATTAACAAATTTTCAACAGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTAATGAAAACTGAAGAGGTAAATCTGTAATCTGG 1560
DB 1501 GCACCTTCAACCAAGGACCATAGATTAATGAAAACTGAAGAGGTAAATCTGTAATCTGG 1560
QY 1561 ATTTAAACCGCGATAAAGGCTATAACCGGCTCGCTGAACTGCGTGAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTTAAACCGCGATAAAGGCTATAACCGGCTCGCTGAACTGCGTGAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
DB 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCGTGTGGCTGAAATCAACCCCGGCAAAAGCGTTTCCAGACAAGCTGTAT 1800
DB 1741 GCTCAATCTGGCGTGTGGCTGAAATCAACCCCGGCAAAAGCGTTTCCAGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACCGCAAGCTGATTTACCGGATCGCTGTT 1860
DB 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACCGCAAGCTGATTTACCGGATCGCTGTT 1860
QY 1861 GAAGGTTTATCGCTGATTTTAAACAAAGATCTGCTGCGCAACCCCGCAAAACCTGGGAA 1920
DB 1861 GAAGGTTTATCGCTGATTTTAAACAAAGATCTGCTGCGCAACCCCGCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
DB 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTTGCTGCTACGGGGGTTATCGGTTCAAG 2040

Db 1981 CTGCAAGAACGTAATTCACTGCGCGCTGATGCTGCTGACGGGGTTATGCTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTAAAGACGTCGGCGTGGATTAACGCTGCGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAAGACGTCGGCGTGGATTAACGCTGCGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAAACCAAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAAACCAAAACACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAATAAGCGGAAACAGCGATGACCATCAACCGCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAATAAGCGGAAACAGCGATGACCATCAACCGCGCTGG 2220
QY 2221 GCATGCTCAACATCGACACCGCAAGTGAATTATGTTGTTAACGCTGCTGCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACCGCAAGTGAATTATGTTGTTAACGCTGCTGCGGACCTTC 2280
QY 2281 AAGGTCACCATCCAAACCGTTGCTGCTGAGCGCAGGTATTAAACGCGCCAGT 2340
Db 2281 AAGGTCACCATCCAAACCGTTGCTGCTGAGCGCAGGTATTAAACGCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGTTTAATAAGACAAACCGCTGGTGGCTGAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGTTTAATAAGACAAACCGCTGGTGGCTGAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGATTTGCGGCCACCACTGGAAGCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTTGCGGCCACCACTGGAAGCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTGATGCGGCTGAGTCTGCGGTGATCAAGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTGATGCGGCTGAGTCTGCGGTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2640
Db 2581 GCCAGCGTCTGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2640
QY 2641 AACACAAACAATAACAATAACAATAACAATAACAATAACAATAACAATAACAATAACAATA 2700
Db 2641 AACACAAACAATAACAATAACAATAACAATAACAATAACAATAACAATAACAATAACAATA 2700

RESULT 14
ID ADO23644 standard; DNA; 7370 BP.
AC ADO23644;
XX
AC
XX
DT 01-JUN-2004 (first entry)
XX
DE DNA encoding MBP-Toxop30MIX3 fusion protein.
XX
XX P30 antigen; Toxo30del3C; Toxo30del2C; ToxoP30 MIX1;
KW MBP-Toxo30del3C(52-300aa); MBP-Toxo30del4C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; gene.
XX
OS Toxoplasma gondii.
OS Synthetic.
XX
XX US2004067239-A1.
XX
XX 08-APR-2004.
XX
XX 02-OCT-2002; 2002US-00263153.
XX
XX 02-OCT-2002; 2002US-00263153.
XX
XX (MAIN/) MAINE G T.
PA

(PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX
PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX
DR WPI; 2004-304563/28.
DR P-P5DB; ADO23645.
XX
PT Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del3C,
PT Toxo30del2C, ToxoP30 MIX1, useful for detecting presence of Igm or Igg
PT antibodies to Toxoplasma gondii.
XX
PS Example 5; Fig 33; 114pp; English.
XX
CC The invention relates to a purified P30 antigen (I) chosen from 3 fully
CC defined Toxo30del3C, Toxo30del2C and ToxoP30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxo30del2C P30 antigen sequence, or comprising an amino acid sequence in
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxo30del3C P30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxo30del3C(52-300aa), MBP-Toxo30del4C(52-294aa) and MBP-Toxo30MIX1
CC fusion proteins (I) is useful for detecting the presence of Igm
CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the Igm antibodies
CC with a composition comprising (I) and detecting the presence of (I)/Igm
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC ToxoP30 fusion protein of the invention.
XX
SQ Sequence 7370 BP; 1900 A; 1857 C; 1950 G; 1663 T; 0 U; 0 Other;
Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGTCGCAAAACCTTTTCGGGTATGCGCATGATAGCCCGGAAAGAGA 60
Db 1 CCGACACCATCGAATGTCGCAAAACCTTTTCGGGTATGCGCATGATAGCCCGGAAAGAGA 60
QY 61 GTCAATTCAGGGTGTGTAATGTGAACACAGTAACCTTATACCATGTGCGCAGAGTATGCCG 120
Db 61 GTCAATTCAGGGTGTGTAATGTGAACACAGTAACCTTATACCATGTGCGCAGAGTATGCCG 120
QY 121 GTGCTCTTTATCAGACCGTTTCCCGGTGTGTAACAGCCAGCCAGCCAGCTTTCTGCGAAAA 180
Db 121 GTGCTCTTTATCAGACCGTTTCCCGGTGTGTAACAGCCAGCCAGCCAGCTTTCTGCGAAAA 180
QY 181 CCGCGGAAAAAGTGAAGCGCGGATGCGGAGCTCAATTACATTTCCCAACCGCGTGGCAC 240
Db 181 CCGCGGAAAAAGTGAAGCGCGGATGCGGAGCTCAATTACATTTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGCGGGGCAAAACAGTCTTGTGCTGATTGGCGTTGCCACCTTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGCGGGGCAAAACAGTCTTGTGCTGATTGGCGTTGCCACCTTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCCGTGCGCAAAATGTGCGCGCGATTAAATCTCGCGCGGATCACTGGTGGTGCACGG 360
Db 301 ACGCGCCGTGCGCAAAATGTGCGCGCGATTAAATCTCGCGCGGATCACTGGTGGTGCACGG 360
QY 361 TGGTGGTGTGATGTTAGAACGAGCGCGTGAAGCCTGTAAAGCGCGGTGACCAATC 420
Db 361 TGGTGGTGTGATGTTAGAACGAGCGCGTGAAGCCTGTAAAGCGCGGTGACCAATC 420
QY 421 TTCTCGGCAACCGGTGATGCGGTGATCAATTAATCCCTGATGATGACGAGATGCCA 480
Db 421 TTCTCGGCAACCGGTGATGCGGTGATCAATTAATCCCTGATGATGACGAGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGGGTTATTTCTTGATGTCCTCACCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGGGTTATTTCTTGATGTCCTCACCAGA 540

QY 541 CACCCATCAACAGTATTATTTCTCCATGAGAGCGTTACCGGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCATGAGAGCGTTACCGGACTGGCGGTGGAGCATC 600
QY 601 TGGTCGCATTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATAAGTTCTGTCTCGG 660
Db 601 TGGTCGCATTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGCTGCGCTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTGCGCTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGACACTGGAGTGCATGTCGGTTTTCAACAACCAATGCAATGCTGA 780
Db 721 CGGAACGGGAAGGACACTGGAGTGCATGTCGGTTTTCAACAACCAATGCAATGCTGA 780
QY 781 ATGAGGGCATCTCCCACTGGGATGCTGGTTCACACGATCAGATGCGCTGGCGCGAA 840
Db 781 ATGAGGGCATCTCTCCCACTGGGATGCTGGTTCACACGATCAGATGCGCTGGCGCGAA 840
QY 841 TGCGCGCATTAACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db 841 TGCGCGCATTAACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACGGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCAATCAACAGGATTTTC 960
Db 901 ACGATACGGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCAATCAACAGGATTTTC 960
QY 961 GCCTCTGGGGCAACACGAGTGGAGCGCTTGCTGCAACTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTCTGGGGCAACACGAGTGGAGCGCTTGCTGCAACTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACCAACCCCTGGCGGCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACCAACCCCTGGCGGCCAATA 1080
QY 1081 CGCAAAACGCTCTCCCGCGGTTGGCGATTCATTAATGCGAGTGGCGACGACAGGTTT 1140
Db 1081 CGCAAAACGCTCTCCCGCGGTTGGCGATTCATTAATGCGAGTGGCGACGACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGGCAGTGCAGCGCAACGAAATTAATGTAGTTAGTCTACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGCAGCGCAACGAAATTAATGTAGTTAGTCTACTCATTAG 1200
QY 1201 GCACAATCTCATGTTTGACAGCTTATCATCGACTGCACGTTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATCTCATGTTTGACAGCTTATCATCGACTGCACGTTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGAGCCATTCGGAAGCTGTGGTATGCTGTGCAAGTTCGTAATCACTGCATTAATTCG 1320
Db 1261 TCAGGAGCCATTCGGAAGCTGTGGTATGCTGTGCAAGTTCGTAATCACTGCATTAATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGTCTGGATATGTTTTTGGCGCGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTGGATATGTTTTTGGCGCGACATCATACGGTT 1380
QY 1381 CTGGCAAAATATCTGAAATGAGCTGTGCAAAATTAATCATCGGCTCGTATAATGTGGA 1440
Db 1381 CTGGCAAAATATCTGAAATGAGCTGTGCAAAATTAATCATCGGCTCGTATAATGTGGA 1440
QY 1441 ATTGTAGCGGATAACAATTTCAACAGAAACAGCCAGTCCGTTAGGTGTTTTCACGA 1500
Db 1441 ATTGTAGCGGATAACAATTTCAACAGAAACAGCCAGTCCGTTAGGTGTTTTCACGA 1500
QY 1501 GCACCTTACCACAAAGGACCATAGCATATGAAAAATCGAAAGAGGTAACCTGTAATCTCG 1560
Db 1501 GCACCTTACCACAAAGGACCATAGCATATGAAAAATCGAAAGAGGTAACCTGTAATCTCG 1560
QY 1561 ATTTAAGCGGATAAAGGCTTAAACGCTCTCGCTGAAGTCCGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTTAAGCGGATAAAGGCTTAAACGCTCTCGCTGAAGTCCGTAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAACCTGGAAGAGAAATTCACAGGTT 1680

Db 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAACCTGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGCAACTGCGCATGCGCCCTGACATTATCTTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGCGCATGCGCCCTGACATTATCTTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCCGACAAAGCGTTCCAGGAACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCCGACAAAGCGTTCCAGGAACAAGCTGTAT 1800
QY 1801 CCGTTTTACTCGGATGCGGTACGTTACAAACGGCAAGCTGATTGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTTACTCGGATGCGGTACGTTACAAACGGCAAGCTGATTGCTTACCCGATCGCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAACGGAAAGTAAGAGCGGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAACGGAAAGTAAGAGCGGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTTCACCTGCGCTGATTGCTGCTGACGGGGTATTGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGCGCTGATTGCTGCTGACGGGGTATTGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTAAAGACATTAAGAACGCTGGGCGTGGATTAACGCTGGCGCGAAACG 2100
Db 2041 TATGAAAAACGCAAGTAAAGACATTAAGAACGCTGGGCGTGGATTAACGCTGGCGCGAAACG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAACAAACACATGAATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAACAAACACATGAATGACAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAAGCTGCGCTTTAATAAGCGCAACAGCAGTACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAAGCTGCGCTTTAATAAGCGCAACAGCAGTACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCAAACATCGACACCAAGTGAATTAATGGTGAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAACATCGACACCAAGTGAATTAATGGTGAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGTCACCAATCAACACCGTTGCTGGCGTCTGAGCGCAGGATTAACGCCGCCAGT 2340
Db 2281 AAGGTCACCAATCAACACCGTTGCTGGCGTCTGAGCGCAGGATTAACGCCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACATATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACATATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCACGATTGCGGCCACCATGGAAGAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTGCGGCCACCATGGAAGAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACCGCCAGACACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACCGCCAGACACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAAACAACCTTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAACAATAAACAACCTTCGGGATCGAGGGAAGGATTCAGAAATTC 2700

ADO23594
ID ADO23594 standard; DNA; 7403 BP.
XX AC
XX ADO23594;
XX AC
XX ADO23594;
XX AC
XX 01-JUL-2004 (first entry)
XX DT
XX DNA encoding MBP-ToxoP30del12(52-311aa) fusion protein.
XX DE
XX P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
XX KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
XX KW Toxoplasma gondii; ds; gene.
XX KW
XX OS Toxoplasma gondii.
XX OS Synthetic.
XX OS
XX PN US2004067239-A1.
XX PD
XX PD 08-APR-2004.
XX PF
XX PF 02-OCT-2002; 2002US-00263153.
XX PR
XX PR 02-OCT-2002; 2002US-00263153.
XX PA (MAIN/) MAINE G T.
XX PA (PATE/) PATEL C B.
XX PA (GINS/) GINSBURG S R.
XX PA (BLIE/) BLIESE T R.
XX PI
XX PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX DR WPI: 2004-304563/28.
XX DR P-PSDB; ADO23595.
XX PT
XX PT Novel purified polypeptide having sequence identity to amino acid
XX PT sequence of p30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
XX PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
XX PT antibodies to Toxoplasma gondii.
XX PS
XX PS Example 2; Fig 11; 114pp; English.
XX CC
XX CC The invention relates to a purified p30 antigen (I) chosen from 3 fully
XX CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
XX CC amino acids added to the C-terminus of the amino acid sequence of
XX CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
XX CC which at least one of the five C-terminal cysteine amino acids of the
XX CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
XX CC with alanine, or comprising the amino acid sequence chosen from MBP-
XX CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
XX CC fusion proteins. (I) is useful for detecting the presence of IgM
XX CC antibodies to Toxoplasma gondii in a test sample, which involves
XX CC contacting the test sample suspected of containing the IgM antibodies
XX CC with a composition comprising (I) and detecting the presence of (I)/IgM
XX CC antibody complexes. The present sequence represents DNA encoding a MBP-
XX CC ToxoP30 fusion protein of the invention.
XX SQ
XX SQ Sequence 7403 BP; 1907 A; 1863 C; 1961 G; 1672 T; 0 U; 0 Other;
Query Match 81.5%; Score 2690.4; DB 12; Length 7403;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGAACACCTTCGCGGTATGGCATATGATAGCCCGCGGAAGAGA 60
DB 1 CCGACACCATCGAATGGTGAACACCTTCGCGGTATGGCATATGATAGCCCGCGGAAGAGA 60
QY 61 GTCAATTCAGGGTGGTGAATGTGAACACCTTCGCGGTATGGCATATGATAGCCCGCGGAAGAGA 120
DB 61 GTCAATTCAGGGTGGTGAATGTGAACACCTTCGCGGTATGGCATATGATAGCCCGCGGAAGAGA 120
QY 121 GTCTCTTATCAGACCGTTTCGCGGTGGTGAACACCTTCGCGGTATGGCATATGATAGCCCGCGGAAGAGA 180
DB 121 GTCTCTTATCAGACCGTTTCGCGGTGGTGAACACCTTCGCGGTATGGCATATGATAGCCCGCGGAAGAGA 180

QY 181 CGCGGAAAAAGTGAAGCGCGCATGGCGAGCTGAATTAATATCCAAACCGCTGGCAC 240
DB 181 CGCGGAAAAAGTGAAGCGCGCATGGCGAGCTGAATTAATATCCAAACCGCTGGCAC 240
QY 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTTCGCAAAATTTGTCGCGCGATTAATATCTCGCGCGGATCAACTGGTGGCGACG 360
DB 301 ACGCGCGTTCGCAAAATTTGTCGCGCGATTAATATCTCGCGCGGATCAACTGGTGGCGACG 360
QY 361 TGGTGGTGTGATGGTAGAAGCGGCGTGAAGCGCTGAAGCGCTGAAGCGCGGTGCACAATC 420
DB 361 TGGTGGTGTGATGGTAGAAGCGGCGTGAAGCGCTGAAGCGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACGCTCAGTGGGCTGATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
DB 421 TTCTCGCGCAACGCTCAGTGGGCTGATCAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
QY 481 TTGCTGTGAAGCTGCTGCACTAAATGTTCCGGCGTTTATTTCTTGATGTTCTCTGACACAGA 540
DB 481 TTGCTGTGAAGCTGCTGCACTAAATGTTCCGGCGTTTATTTCTTGATGTTCTCTGACACAGA 540
QY 541 CACCCATCAACAGTATTTATTTCTCCATGAAGCGGTACGCGACTGGCGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTTATTTCTCCATGAAGCGGTACGCGACTGGCGTGGAGCATC 600
QY 601 TGGTTCGCAATTTGGCTCACAGCAATCGCGCTGTTAGCGGCGCCATTAAGTTCCTCTCGG 660
DB 601 TGGTTCGCAATTTGGCTCACAGCAATCGCGCTGTTAGCGGCGCCATTAAGTTCCTCTCGG 660
QY 661 CGGCTCTGCTTGGCTGGCTGCAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
DB 661 CGGCTCTGCTTGGCTGGCTGCAATAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 720
QY 721 CGGAACGGGAAGCGAGCTGGAGTGGCATGTCGGTTTTTCAACAAACCATGCAAAATCTGA 780
DB 721 CGGAACGGGAAGCGAGCTGGAGTGGCATGTCGGTTTTTCAACAAACCATGCAAAATCTGA 780
QY 781 ATGAGGCGCATCGTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
DB 781 ATGAGGCGCATCGTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 841 TGGCGCGCATTAACGAGTTCGGGCTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGG 900
DB 841 TGGCGCGCATTAACGAGTTCGGGCTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGG 900
QY 901 AGCATACCGAAGACAGCTCATGTTATATATCCCGCGTTTACCAACCATCAACAGGATTTTC 960
DB 901 AGCATACCGAAGACAGCTCATGTTATATATCCCGCGTTTACCAACCATCAACAGGATTTTC 960
QY 961 GCTGCTGGGGCAAAACAGCGTGGACCGCTTCTGCACTCTCTCAGGGCCAGGGGGTGA 1020
DB 961 GCTGCTGGGGCAAAACAGCGTGGACCGCTTCTGCACTCTCTCAGGGCCAGGGGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCGCGTCTCACTGGTGAAGAAAAACCAACCTTGGCGCGCAATA 1080
DB 1021 AGGGCAATCAGCTGTTGCGCGTCTCACTGGTGAAGAAAAACCAACCTTGGCGCGCAATA 1080
QY 1081 CGCAAAACCGCTTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAGCAGAGTTT 1140
DB 1081 CGCAAAACCGCTTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAGCAGAGTTT 1140
QY 1141 CCGGACTGGAAGCGGGCAGTGGCGCAACCGCAATTAATGAGCTGGCAGCAGAGTTT 1200
DB 1141 CCGGACTGGAAGCGGGCAGTGGCGCAACCGCAATTAATGAGCTGGCAGCAGAGTTT 1200
QY 1201 GCACAAATTTCTCATGTTTGAACAGCTTATCATCGACTGACCGGTGACCAATGCTTCTGCGG 1260
DB 1201 GCACAAATTTCTCATGTTTGAACAGCTTATCATCGACTGACCGGTGACCAATGCTTCTGCGG 1260

QY 1261 TCAGCAGCCATCGAAGCTGTGGTATGGCTGTGCAGTCTGTAATCACTCATTAATTCG 1320
Db 1261 TCAGCAGCCATCGAAGCTGTGGTATGGCTGTGCAGTCTGTAATCACTCATTAATTCG 1320
QY 1321 TGTGCTCAAGCGCAGCTCCCGTTCGTGATAATGTTTTTGGCGCGACATCATAAACGGTT 1380
Db 1321 TGTGCTCAAGCGCAGCTCCCGTTCGTGATAATGTTTTTGGCGCGACATCATAAACGGTT 1380
QY 1381 CTGGCAAAATATCTGAAATAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGTGA 1440
Db 1381 CTGGCAAAATATCTGAAATAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGTGA 1440
QY 1441 ATTGTGAGGGATACAAATTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCAGA 1500
Db 1441 ATTGTGAGGGATACAAATTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCAGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAAACTGAAAGAGTAAACTGGTAACTCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGCATATGAAATCGAAGAGTAAACTGGTAACTCTGG 1560
QY 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCAATGCGGATGGCCCTGACATTTATCTCGGCACACAGCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAATGCGGATGGCCCTGACATTTATCTCTGGGCACACAGCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTGGCTGAATCAACCCCGGACAAAGCGTTCCAGSACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTGGCTGAATCAACCCCGGACAAAGCGTTCCAGSACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGGCGTAGCTTTACACGGCAAGCTGATTTGCTTACCGCATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGGCGTAGCTTTACACGGCAAGCTGATTTGCTTACCGCATCGCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTCGGAACCCCGCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTCGGAACCCCGCAAAACCTGGGAA 1920
QY 1921 GAGATCCGCGGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGCGGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTTGCTGACGGGGGTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTTGCTGACGGGGGTTATCGGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGGCGTGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGGCGTGATTAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTGTGACCTGATTAATAAACAACACATGAATGCGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTGTGACCTGATTAATAAACAACACATGAATGCGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCAACTCGACACCGCAAGTGAATTAATGTTAAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAACTCGACACCGCAAGTGAATTAATGTTAAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCAAAACCGTTTGGCTGCTGAGCGCAGGTATTAAACGGCCAGT 2340
Db 2281 AAGGGTCAACCATCAAAACCGTTTGGCTGCTGAGCGCAGGTATTAAACGGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACATATCTGCTGATGAAGTCTG 2400

Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACATATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAAGACAAAACCGCTGGGTCCCGTAGCGCTGAAAGTCTTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAAACCGCTGGGTCCCGTAGCGCTGAAAGTCTTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCACAGTATTTGCCGCCACCATGGGAAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACAGTATTTGCCGCCACCATGGGAAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTTGGTATGCGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTTGGTATGCGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCCGACACTTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCCGACACTTAATTCGAGCTCG 2640
QY 2641 AACCAACAAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACCAACAAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700

Search completed: August 1, 2005, 04:28:37
Job time : 1101.6 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 20:22:50 ; Search time 9276 Seconds
(without alignments)
17238.272 Million cell updates/sec

Title: US-09-765-555B-15
Perfect score: 3300
Sequence: 1 ccgacacatcgaaatgtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*

2: gb_hcg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_ats.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3300	100.0	3300	6	AX202426	Sequence
2	3300	100.0	3300	6	AX202427	Sequence
3	3226.4	97.8	3300	6	AX202430	Sequence
4	3215.2	97.4	3300	6	AX202429	Sequence
5	3207.2	97.2	3300	6	AX202428	Sequence
6	2705.8	82.0	7373	6	AX284144	Sequence
7	2705.8	82.0	8101	6	AX172306	Sequence
8	2701	81.8	6648	6	AX377531	Sequence
9	2700	81.8	7475	6	AX1965	Sequence
10	2700	81.8	7475	6	AR031992	Sequence
11	2700	81.8	7475	6	AR207294	Sequence
12	2698	81.8	9191	6	AX377532	Sequence
13	2676.2	81.1	6724	12	AF097412	Expressio
14	2636.8	79.9	6806	6	AX378208	Sequence
15	2631	79.7	6706	12	AF031088	Shuttle v
16	2598.2	78.7	6748	12	AF031813	Expressio
17	1609.4	48.8	5558	6	CQ794769	Sequence
18	1474	44.7	4700	6	CQ846980	Sequence
19	1462.4	44.3	4921	6	AX768201	Sequence

	20	1462.4	44.3	4935	6	AX768200	Sequence
	21	1462.4	44.3	4945	6	AX768202	Sequence
	22	1462.4	44.3	4951	6	AX768203	Sequence
C	23	1432.4	43.4	6755	12	AF140577	Integrati
	24	1417.6	43.0	1511	6	AX431576	Sequence
	25	1417.6	43.0	4297	12	AY219685	Expressio
	26	1417.6	43.0	4359	12	AY219686	Expressio
	27	1417.6	43.0	4419	12	AY219688	Expressio
	28	1417.6	43.0	4855	12	AY219689	Expressio
	29	1417.6	43.0	4968	12	AY219687	Expressio
	30	1417.6	43.0	6956	12	AY219682	Shuttle e
	31	1417.6	43.0	7018	12	AY219683	Shuttle e
	32	1417.6	43.0	7509	12	AY219684	Shuttle e
	33	1400.8	42.4	4566	12	AF177933	Cloning v
	34	1400.8	42.4	4635	12	AF177932	Cloning v
	35	1400.8	42.4	6399	12	AY302759	Expressio
	36	1400.8	42.4	6587	12	AY289542	Expressio
	37	1362	41.3	5201	6	AR493834	Sequence
	38	1362	41.3	5201	6	AR493835	Sequence
	39	1329.2	40.3	4806	12	AF050464	Expressio
C	40	1238.4	37.5	4803	12	AY243506	Cloning v
C	41	1238.4	37.5	5024	6	CQ846977	Sequence
	42	1237.2	37.5	1922	6	AX478528	Sequence
	43	1204	36.5	4506	12	AB116367	Expressio
	44	1204	36.5	4557	6	E13740	DNA sequenc
	45	1204	36.5	4557	6	AR373251	Sequence

ALIGNMENTS

RESULT 1
AX202426
LOCUS AX202426 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 14 from Patent WO0152620.
ACCESSION AX202426
VERSION AX202426.1 GI:15392173
KEYWORDS synthetic construct
SOURCE other sequences; artificial sequences.
ORGANISM
REFERENCE 1
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 14 26-JUL-2001;
The Scripps Research Institute (US) ; SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES
source Location/Qualifiers
1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-m1 and zinc finger protein ZFPm1"

ORIGIN	Query Match	100.0%;	Score 3300;	DB 6;	Length 3300;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 3300;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	CGACACCATCGAATGGTGC	AAACCTTTCCGCGGTATG	GCATGATGATAGCCCGGA	AGAGA 60
Db	1	CGACACCATCGAATGGTGC	AAACCTTTCCGCGGTATG	GCATGATGATAGCCCGGA	AGAGA 60
QY	61	GTCAATTCAGGTGGTGAAT	GTGAACCACTTATACGAT	TCCGATGATGATGATGCG	120
Db	61	GTCAATTCAGGTGGTGAAT	GTGAACCACTTATACGAT	TCCGATGATGATGATGCG	120
QY	121	GTGTCTCTTATCAGACCG	GTTCCTCCCGGTGGTGA	ACCCAGCCAGCCACGCTT	CTCGGAAAA 180
Db	121	GTGTCTCTTATCAGACCG	GTTCCTCCCGGTGGTGA	ACCCAGCCAGCCACGCTT	CTCGGAAAA 180
QY	181	CGCGGAAAAAGTGGAAG	CGCGGATGCGGAGCTGA	ATTACATTCCTCCACCG	CGTGGCAC 240

181 CGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTACATTCCTCCAAACCGCTGGCAC 240
241 ACAAACCTGGCGGGAAAAAGTGTGCTGTTGATGGCGTTGGCACTCCAGTCTGGCCCTGC 300
241 ACAAACCTGGCGGGAAAAAGTGTGCTGTTGATGGCGTTGGCACTCCAGTCTGGCCCTGC 300
301 AGCGCCGCTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGATCAATCTGGTGGCGAGG 360
301 AGCGCCGCTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGATCAATCTGGTGGCGAGG 360
361 TGGTGGTTCGATGTGAAGCAAGCGCGCTCGAAGCGCTGTAAGCGCGGTGCAACAATC 420
361 TGGTGGTTCGATGTGAAGCAAGCGCGCTCGAAGCGCTGTAAGCGCGGTGCAACAATC 420
421 TTCTCGCCCAACGCTCAGTGGCGGTGATCAATTAATCTCGCTGGATGACCAAGATGCCA 480
421 TTCTCGCCCAACGCTCAGTGGCGGTGATCAATTAATCTCGCTGGATGACCAAGATGCCA 480
481 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCCGGCGTTAATTTCTTGATGTTCTGACCA 540
481 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCCGGCGTTAATTTCTTGATGTTCTGACCA 540
541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGACTGGCGGTGGAGCATC 600
541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGACTGGCGGTGGAGCATC 600
601 TGGTGCATTCGGTCCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCCTCTCGG 660
601 TGGTGCATTCGGTCCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCCTCTCGG 660
661 CGCGTCTCGCTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
661 CGCGTCTCGCTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
721 CGGAACGGGAAGGCGATGGAGTGCATGTCGGTTCATCAAAACCAATCAAAATGCTGA 780
721 CGGAACGGGAAGGCGATGGAGTGCATGTCGGTTCATCAAAACCAATCAAAATGCTGA 780
781 ATGAGGGCATCGTTCACCTCCGATGCTGTTGCGCAACGATCAGATGGCGTGGCCGCAA 840
781 ATGAGGGCATCGTTCACCTCCGATGCTGTTGCGCAACGATCAGATGGCGTGGCCGCAA 840
841 TGCGCGCATTACCGAGTCCGGGCTGGCGTGTGCGGATATCTCGGTAGTGGATACG 900
841 TGCGCGCATTACCGAGTCCGGGCTGGCGTGTGCGGATATCTCGGTAGTGGATACG 900
901 ACGATACGGAAGACAGCTCATGTTATATCCGCGGTTAAACCATCAACAGGATTTTC 960
901 ACGATACGGAAGACAGCTCATGTTATATCCGCGGTTAAACCATCAACAGGATTTTC 960
961 GCCTGTGGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
961 GCCTGTGGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
1021 AGGCAATCAGCTGTTCGCCGTCTCACTGGTGAAGAAAAACAACCCCTGGCGCCCAATA 1080
1021 AGGCAATCAGCTGTTCGCCGTCTCACTGGTGAAGAAAAACAACCCCTGGCGCCCAATA 1080
1081 CGCAACCGCTCTCCCGCGGTTGGCGATTCATTAATGAGTTCAGCTCACTCAATTAG 1140
1081 CGCAACCGCTCTCCCGCGGTTGGCGATTCATTAATGAGTTCAGCTCACTCAATTAG 1140
1141 CCCGACTGGAAGCGGCGAGTGAAGCAACGCAATTAATGTGAGTTAGCTCACTCAATTAG 1200
1141 CCCGACTGGAAGCGGCGAGTGAAGCAACGCAATTAATGTGAGTTAGCTCACTCAATTAG 1200
1201 GCACAAATCTCATGTTTGACAGCTTATCATGATGCAACGCGTGAACAAATGCTTCGGG 1260
1201 GCACAAATCTCATGTTTGACAGCTTATCATGATGCAACGCGTGAACAAATGCTTCGGG 1260
1261 TCAGGCAGCCATCGAAGCTGTGGTATGGCTGTGAGGTGCTGAATCACTGCAATAATTCG 1320

1261 TCAGGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCTGTAAATCACTGCATAATTCG 1320
1321 TGTGCTCAAGCGCAGCTCCCGTTCTGGATAATGTTTTTTTGGCGCGACATCATACGGTT 1380
1321 TGTGCTCAAGCGCAGCTCCCGTTCTGGATAATGTTTTTTTGGCGCGACATCATACGGTT 1380
1381 CTGCAAAATTTCTGAAATGAGCTGTGGACAAATTAATCATCGCTCGTATAATGTGGGA 1440
1381 CTGCAAAATTTCTGAAATGAGCTGTGGACAAATTAATCATCGCTCGTATAATGTGGGA 1440
1441 ATTCTGAGCGGATAAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAAAGTGTTCACGA 1500
1441 ATTCTGAGCGGATAAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAAAGTGTTCACGA 1500
1501 GCATTTCCACCAACAAAGGACCATAGATTATGAAATCTGAAGAGGTAACCTGGTAAATCTGG 1560
1501 GCATTTCCACCAACAAAGGACCATAGATTATGAAATCTGAAGAGGTAACCTGGTAAATCTGG 1560
1561 ATTAAACGGCGATTAAGGCTATACCGTCTCGCTGAAGTCCGGTAAGAAATTCGAGAAAGAT 1620
1561 ATTAAACGGCGATTAAGGCTATACCGTCTCGCTGAAGTCCGGTAAGAAATTCGAGAAAGAT 1620
1621 ACCGGAATTAAGTCAACCGTTGAGCATCCCGATTAACCTGGAAGAGAAATTCGCCACAGTT 1680
1621 ACCGGAATTAAGTCAACCGTTGAGCATCCCGATTAACCTGGAAGAGAAATTCGCCACAGTT 1680
1681 GCGGCAACTGGCGATGGCCCTGACATTTCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
1681 GCGGCAACTGGCGATGGCCCTGACATTTCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
1741 GCTCAATCTGGCGTGTGGCTGAATCAACCGCAAGCTGATTCCTACCGGATCGCTGTT 1800
1741 GCTCAATCTGGCGTGTGGCTGAATCAACCGCAAGCTGATTCCTACCGGATCGCTGTT 1800
1801 CCGTTTACCTGGGATGCGGTACCGTTACACCGCAAGCTGATTCCTACCGGATCGCTGTT 1860
1801 CCGTTTACCTGGGATGCGGTACCGTTACACCGCAAGCTGATTCCTACCGGATCGCTGTT 1860
1861 GAAGCGTTATCGCTGATTTATACAAAGATCTGCTGCCGAACCCGCGCAAAACCTGGAA 1920
1861 GAAGCGTTATCGCTGATTTATACAAAGATCTGCTGCCGAACCCGCGCAAAACCTGGAA 1920
1921 GAGATCCCGCGCTGATTAAGAACTGAAAGCGAAGGTAAAGCGCGCTGATGTTCAAC 1980
1921 GAGATCCCGCGCTGATTAAGAACTGAAAGCGAAGGTAAAGCGCGCTGATGTTCAAC 1980
1981 CTGCAAGAACCGTATCTTCACTGGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
1981 CTGCAAGAACCGTATCTTCACTGGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
2041 TATGAAACGGCAAGTACGACATTAAGACGTTGGGCTGGATAACGCTGGCGGAAGCG 2100
2041 TATGAAACGGCAAGTACGACATTAAGACGTTGGGCTGGATAACGCTGGCGGAAGCG 2100
2101 GGTCTGACCTTCTGCTGTTGACCTGATTAATAAACAACATGATGACACACCGGATTAC 2160
2101 GGTCTGACCTTCTGCTGTTGACCTGATTAATAAACAACATGATGACACACCGGATTAC 2160
2161 TCCATCGCAAGAGCTGCCCTTTAATAAGCGGAAACAGCGATGACCATCAACGCGCGTGG 2220
2161 TCCATCGCAAGAGCTGCCCTTTAATAAGCGGAAACAGCGATGACCATCAACGCGCGTGG 2220
2221 GATGCTCCAAACATCGACACGCAAGGTAATGTTGCTGCTGAGCGCAGGTATTAAACCGCGCAGT 2280
2221 GATGCTCCAAACATCGACACGCAAGGTAATGTTGCTGCTGAGCGCAGGTATTAAACCGCGCAGT 2280
2281 AAGGGTCAACCATCCAAACCGTTTGGTGGCTGCTGAGCGCAGGTATTAAACCGCGCAGT 2340
2281 AAGGGTCAACCATCCAAACCGTTTGGTGGCTGCTGAGCGCAGGTATTAAACCGCGCAGT 2340
2341 CCGAACAAAGAGCTGGCAAAAGATTCCTCGAATACTATCTGCTGACTGATGAAGTCTG 2400
2341 CCGAACAAAGAGCTGGCAAAAGATTCCTCGAATACTATCTGCTGACTGATGAAGTCTG 2400

QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTCCGTAGCGCTGAAGTCTTACGAGGAAG 2460
D 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTCCGTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGCTATTGCGCCACCACTGGGAAACGCCACAGAAAGGTGAATCATG 2520
D 2461 TTGGCGAAAGATCCACGCTATTGCGCCACCACTGGGAAACGCCACAGAAAGGTGAATCATG 2520
QY 2521 CCGAATCCCGCAGATGTCGCTTTCTGGTATGCCGTGCTGCTGATCAACGCC 2580
D 2521 CCGAATCCCGCAGATGTCGCTTTCTGGTATGCCGTGCTGCTGATCAACGCC 2580
QY 2581 GCCAGCGTGTGACACTGTGATGAAGCCCTGAAAGAGCGGCAGACTAAATTCGAGCTCG 2640
D 2581 GCCAGCGTGTGACACTGTGATGAAGCCCTGAAAGAGCGGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAAACAACAACCACTCGGGATCGAGGGAAGATTTCAGAAATTC 2700
D 2641 AACAAACAACAATAAACAACAACCACTCGGGATCGAGGGAAGATTTCAGAAATTC 2700
QY 2701 GGATCCTCTTCTGTGGCCAGCGCCCTCGAGCCCGGGGAGAACCTTATCTTGT 2760
D 2701 GGATCCTCTTCTGTGGCCAGCGCCCTCGAGCCCGGGGAGAACCTTATCTTGT 2760
QY 2761 CCGGAATGTGTAAGTCTTCTCTCAGAGCTCTACCTGCTGCGGCACAGCGGTACCCAC 2820
D 2761 CCGGAATGTGTAAGTCTTCTCTCAGAGCTCTACCTGCTGCGGCACAGCGGTACCCAC 2820
QY 2821 ACGGTTGAAACCACTGATTAATGCCAGAGTGGCGCAAACTTTTAGCCAGTCCAGCAAC 2880
D 2821 ACGGTTGAAACCACTGATTAATGCCAGAGTGGCGCAAACTTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGGTGCCTATCAACGCACTCATACTGCGGAGAACCATACAAATGTCCAGAAATGTGCG 2940
D 2881 CTGGTGCCTATCAACGCACTCATACTGCGGAGAACCATACAAATGTCCAGAAATGTGCG 2940
QY 2941 AAGTCTTCTCTCGGTCTGACAACTCTGTCGGGACCAACAGTACTCACACCGGGGAGAAG 3000
D 2941 AAGTCTTCTCTCGGTCTGACAACTCTGTCGGGACCAACAGTACTCACACCGGGGAGAAG 3000
QY 3001 CCTATGCTTGTCCGGAATGTGGTAAGTCTTTCAGCCGAGGATAACTGCTGGTGGCCAC 3060
D 3001 CCTATGCTTGTCCGGAATGTGGTAAGTCTTTCAGCCGAGGATAACTGCTGGTGGCCAC 3060
QY 3061 CAGCGTACCCACACCGGTGAAACCGGTATAATGCCAGAGTGGCGCAAACTTTTATAGC 3120
D 3061 CAGCGTACCCACACCGGTGAAACCGGTATAATGCCAGAGTGGCGCAAACTTTTATAGC 3120
QY 3121 CAGCGCGGCCACTGGCGAGCCATCAACGCACTCATACTGCGGAGAACCATACAAATGT 3180
D 3121 CAGCGCGGCCACTGGCGAGCCATCAACGCACTCATACTGCGGAGAACCATACAAATGT 3180
QY 3181 CCAGAAATGGCAAGTCTTCTCTCGGTGCAAACTCTGCTGCAAACTCTGCTGCAAACTCT 3240
D 3181 CCAGAAATGGCAAGTCTTCTCTCGGTGCAAACTCTGCTGCAAACTCTGCTGCAAACTCT 3240
QY 3241 ACCGTTAAAAAATAGTGGCCAGCGCCGAGTACCGGTACGAGTTCGCGACTACGCT 3300
D 3241 ACCGTTAAAAAATAGTGGCCAGCGCCGAGTACCGGTACGAGTTCGCGACTACGCT 3300

RESULT 2
LOCUS AX202427 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 15 from Patent WO0152620.
ACCESSION AX202427
VERSION AX202427.1 GI:15392175
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1

AUTHORS Barbas, C.P., Stege, J.T., Guan, X. and Dalmia, B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 15 26-JUL-2001;
The Scripps Research Institute (US); SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES Location/Qualifiers
source 1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-m2 and zinc finger protein
ZFPm2"
ORIGIN
Query Match 100.0%; Score 3300; DB 6; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGAACACCTTTCGGGTATGGCATGTAGCGCCGGAAGAGA 60
D 1 CCGACACCATCGAATGGTGAACACCTTTCGGGTATGGCATGTAGCGCCGGAAGAGA 60
QY 61 GTCAATTTCAGGCTGGTGAATGTGAACCAAGTAAAGTTATAGATGTCGACAGATATGCG 120
D 61 GTCAATTTCAGGCTGGTGAATGTGAACCAAGTAAAGTTATAGATGTCGACAGATATGCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCGCGTGTGAACAGGCGAGCCACGCTTTCGCAAAA 180
D 121 GTGCTCTTATCAGACCGTTTCCGCGTGTGAACAGGCGAGCCACGCTTTCGCAAAA 180
QY 181 CGCGGAAAAAGTGGAGCGCGGATGGCGGAGCTGAATTAATCCCAACCGCTGGCAC 240
D 181 CGCGGAAAAAGTGGAGCGCGGATGGCGGAGCTGAATTAATCCCAACCGCTGGCAC 240
QY 241 AACAACTGCGCGGCAACAGTCTGCTGATTTGGGTTGGCCACCTCCAGTCTGGGCCCTGC 300
D 241 AACAACTGCGCGGCAACAGTCTGCTGATTTGGGTTGGCCACCTCCAGTCTGGGCCCTGC 300
QY 301 ACGCGCGTCCGCAATTTGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCGACG 360
D 301 ACGCGCGTCCGCAATTTGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCGACG 360
QY 361 TGGTGGTTCGATGGTAGAACGAGCGCGTGAAGCTGTAAAGCGGGTGGCAATC 420
D 361 TGGTGGTTCGATGGTAGAACGAGCGCGTGAAGCTGTAAAGCGGGTGGCAATC 420
QY 421 TTCTCGCCCAACGCTCAGTGGCTGATTAATTAATCTATCCGCTGATGACAGGATGCCA 480
D 421 TTCTCGCCCAACGCTCAGTGGCTGATTAATTAATCTATCCGCTGATGACAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTGATGTTCTGACCA 540
D 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTGATGTTCTGACCA 540
QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACGGGCTGGGCTGGGAGCATC 600
D 541 CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACGGGCTGGGCTGGGAGCATC 600
QY 601 TGGTCCGATTTGGTTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCCTG 660
D 601 TGGTCCGATTTGGTTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCCTG 660
QY 661 CGCGTCTGCGTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
D 661 CGCGTCTGCGTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGCGACTGGAGTGCATGTCGGTTTTCACAAACCATGCAATGCTGA 780
D 721 CGGAACGGGAAGCGACTGGAGTGCATGTCGGTTTTCACAAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTCGATGCTGTTGCCAAGATCAGATGCGCTGGCGCA 840
D 781 ATGAGGGCATGTTCCCACTCGATGCTGTTGCCAAGATCAGATGCGCTGGCGCA 840

QY	841	TGCGGCCATTACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG	900
DB	841	TGCGGCCATTACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACCAACCAAGGATTTTC	960
DB	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACCAACCAAGGATTTTC	960
QY	961	GCCTGTGGGCAACACAGCGTGACCGCTTCTGCGAATCTCTCAGGCGCAGCGGTGA	1020
DB	961	GCCTGTGGGCAACACAGCGTGACCGCTTCTGCGAATCTCTCAGGCGCAGCGGTGA	1020
QY	1021	AGGCAATCAGCTGTGTCGCGCTCTCACTGTGGAAGAAACCAACCCCTGCGGCCAATA	1080
DB	1021	AGGCAATCAGCTGTGTCGCGCTCTCACTGTGGAAGAAACCAACCCCTGCGGCCAATA	1080
QY	1081	CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCATTAATGAGTGGCACGAGGTTT	1140
DB	1081	CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCATTAATGAGTGGCACGAGGTTT	1140
QY	1141	CCCGACTGGAAGCGGCGAGTGAGCGAACGCAATTAATGTGAGTTAGCTCACTCATTAG	1200
DB	1141	CCCGACTGGAAGCGGCGAGTGAGCGAACGCAATTAATGTGAGTTAGCTCACTCATTAG	1200
QY	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACGCGTGCAACCAATGCTTCGGCG	1260
DB	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACGCGTGCAACCAATGCTTCGGCG	1260
QY	1261	TCAGGACGCATCGAAGCTGTGATGCTGTGAGTGTGAGTGTGTAATCACTGCATAATCG	1320
DB	1261	TCAGGACGCATCGAAGCTGTGATGCTGTGAGTGTGAGTGTGTAATCACTGCATAATCG	1320
QY	1321	TGTCGCTCAAGCGCACCTCCCGTTCTGATAATGTTTTGCGCGGACATCAACCGTT	1380
DB	1321	TGTCGCTCAAGCGCACCTCCCGTTCTGATAATGTTTTGCGCGGACATCAACCGTT	1380
QY	1381	CTGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA	1440
DB	1381	CTGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA	1440
QY	1441	ATTGTGAGCGGATAACAAATTTACACAGGAAACAGCCAGTCCGTTTGGTGTTCACGA	1500
DB	1441	ATTGTGAGCGGATAACAAATTTACACAGGAAACAGCCAGTCCGTTTGGTGTTCACGA	1500
QY	1501	GCACCTTCAACCAAGGACCATAGATTATGAAAATGAAAGGTAACCTGTAATCTGG	1560
DB	1501	GCACCTTCAACCAAGGACCATAGATTATGAAAATGAAAGGTAACCTGTAATCTGG	1560
QY	1561	ATTAAACCGCGATAAAGCTATACCGGTCTCGCTGAAGTTCGGTAAAGAAATTCGAGAAAGAT	1620
DB	1561	ATTAAACCGCGATAAAGCTATACCGGTCTCGCTGAAGTTCGGTAAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAAGTTCAGGTTGAGCTCCGGATAACTGGNAGAAATTCACAGGTT	1680
DB	1621	ACCGGAATTAAGTTCAGGTTGAGCTCCGGATAACTGGNAGAAATTCACAGGTT	1680
QY	1681	CGGCAACTGCGCATGCGCTTGAATTTCTTCTGGGCAACACGACCGCTTGGTGGCTAC	1740
DB	1681	CGGCAACTGCGCATGCGCTTGAATTTCTTCTGGGCAACACGACCGCTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGCGCTGTTGGTGAATCAACCCCGGACAAAGGTTTCCAGGCAAGCTGTAT	1800
DB	1741	GCTCAATCTGCGCTGTTGGTGAATCAACCCCGGACAAAGGTTTCCAGGCAAGCTGTAT	1800
QY	1801	CCGTTTACCTGGGATCCGTCAGTTTACACGGCAAGCTGATTCGTTACCGATCGCTGTT	1860
DB	1801	CCGTTTACCTGGGATCCGTCAGTTTACACGGCAAGCTGATTCGTTTACCGATCGCTGTT	1860
QY	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGTCGCGAACC CGCCAAAACCTCGGAA	1920
DB	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGTCGCGAACC CGCCAAAACCTCGGAA	1920

QY	1921	GAGATCCGCGCTGGATATAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
DB	1921	GAGATCCGCGCTGGATATAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTCCTGCTGACCGGGGTTATCGCTTCAAG	2040
DB	1981	CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTCCTGCTGACCGGGGTTATCGCTTCAAG	2040
QY	2041	TATGAAACCGGCAAGTACGACATTAAGAGCTGGCGTGGATTAACGCTGGCGAAAGCG	2100
DB	2041	TATGAAACCGGCAAGTACGACATTAAGAGCTGGCGTGGATTAACGCTGGCGAAAGCG	2100
QY	2101	GGTCTGACCTTCTGCTGTTGACCTGATTTAAACCAACACATGATGACAGACCGGATTAC	2160
DB	2101	GGTCTGACCTTCTGCTGTTGACCTGATTTAAACCAACACATGATGACAGACCGGATTAC	2160
QY	2161	TCCATCGCAGAAGCTGCTTTTAAATAAGGCGAAACAGCGATGACCATCAACGCGCCGTGG	2220
DB	2161	TCCATCGCAGAAGCTGCTTTTAAATAAGGCGAAACAGCGATGACCATCAACGCGCCGTGG	2220
QY	2221	GCATGTCCAAACATCGACACACAGCAAAAGTGAATTAATGTTAAACGCTGACGACCTTC	2280
DB	2221	GCATGTCCAAACATCGACACACAGCAAAAGTGAATTAATGTTAAACGCTGACGACCTTC	2280
QY	2281	AAGGTCACCAATCCAAACCGTTTGGCGTCTGAGCGCAGGTATTTAACGCGCCAGT	2340
DB	2281	AAGGTCACCAATCCAAACCGTTTGGCGTCTGAGCGCAGGTATTTAACGCGCCAGT	2340
QY	2341	CCGAAACAAAGAGCTGCGCAAAAGGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG	2400
DB	2341	CCGAAACAAAGAGCTGCGCAAAAGGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG	2400
QY	2401	GAAGCGGTTTAAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTACGAGGAAGAG	2460
DB	2401	GAAGCGGTTTAAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTACGAGGAAGAG	2460
QY	2461	TTGGCGAAAGATCCACGTATTGCGCGCACCATCGAAACCGCCAGAAAGGTGAATCATG	2520
DB	2461	TTGGCGAAAGATCCACGTATTGCGCGCACCATCGAAACCGCCAGAAAGGTGAATCATG	2520
QY	2521	CCGAAACATCCGCGAGATGTCGCGTTTCTGATATGCGGTGCGTACTGCGGTGATCAACGCC	2580
DB	2521	CCGAAACATCCGCGAGATGTCGCGTTTCTGATATGCGGTGCGTACTGCGGTGATCAACGCC	2580
QY	2581	GCAGCGGTCGTGAGACTGTCGATGAAGCCCTGAAAGACGCGAGACTAATTCGAGCTCG	2640
DB	2581	GCAGCGGTCGTGAGACTGTCGATGAAGCCCTGAAAGACGCGAGACTAATTCGAGCTCG	2640
QY	2641	AACAAACAAACAAATTAACAACTTAAACCAACCTCGGATCGAGGGAAGGATTTTCAAGATTC	2700
DB	2641	AACAAACAAACAAATTAACAACTTAAACCAACCTCGGATCGAGGGAAGGATTTTCAAGATTC	2700
QY	2701	GGATCTCTTCTGTCGCGCGCGGCTCGAGCGCGGGGAGAGCCCTATGCTTGT	2760
DB	2701	GGATCTCTTCTGTCGCGCGCGGCTCGAGCGCGGGGAGAGCCCTATGCTTGT	2760
QY	2761	CCGGAATGTTGTAAGTCTTCTCTCAGAGCTCTCACTGCTGCGCCACAGCGTACCAC	2820
DB	2761	CCGGAATGTTGTAAGTCTTCTCTCAGAGCTCTCACTGCTGCGCCACAGCGTACCAC	2820
QY	2821	ACGGGTGAAACCGGTATATAATGCCAGAGTGGCGAAATCTTTTACGCCAGTCCAGCAAC	2880
DB	2821	ACGGGTGAAACCGGTATATAATGCCAGAGTGGCGAAATCTTTTACGCCAGTCCAGCAAC	2880
QY	2881	CTGGTGGCCATCAACCGACTCATATCTGGGAGAACCCATAAATGTCAGAAATGTGGC	2940
DB	2881	CTGGTGGCCATCAACCGACTCATATCTGGGAGAACCCATAAATGTCAGAAATGTGGC	2940
QY	2941	AAGTCTTTCTCTCGGTCTGACAACTCTGTCGCGCACCAACGCTACTCACACCGGGGAGAG	3000
DB	2941	AAGTCTTTCTCTCGGTCTGACAACTCTGTCGCGCACCAACGCTACTCACACCGGGGAGAG	3000
QY	3001	CCCTATGCTGTCCGGAATGTGGTAAAGTCTTTCAGCGCGCAGCGATTAACCTGTGCGCCAC	3060

Db 3001 CCCTATGCTTGTCCGGAATGTGGTAAGTCCCTTCAGCCGAGGGAATAACCTGGTGGCCAC 3060
Qy 3061 CAGCGTACCCACACACGGGTGAAAAACCGTATATAATGCCAGAGTGGCGAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACACGGGTGAAAAACCGTATATAATGCCAGAGTGGCGAAATCTTTTAGC 3120
Qy 3121 CAGCGCGGCACCTGGCCAGCCCATCAACGCACTATCTACTGGGAGAACCATACAAATGT 3180
Db 3121 CAGCGCGGCACCTGGCCAGCCCATCAACGCACTATCTACTGGGAGAACCATACAAATGT 3180
Qy 3181 CCAGAAATGTGGCAAGTCTTTCTCTCGGTCTGACAATCTCGTCGGCACCACAGTACTCAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTTCTCTCGGTCTGACAATCTCGTCGGCACCACAGTACTCAC 3240
Qy 3241 ACCCGTAAAAAACTAGTGGCCAGCGCCGCAAGTACCGGTACGAGCTTCCGCACTACGCT 3300
Db 3241 ACCCGTAAAAAACTAGTGGCCAGCGCCGCAAGTACCGGTACGAGCTTCCGCACTACGCT 3300

RESULT 3
AX202430
LOCUS AX202430 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 18 from Patent WO0152620.
ACCESSION AX202430
VERSION AX202430.1 GI:15392178
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 18 26-JUL-2001;
The Scripps Research Institute (US); SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)

FEATURES
source
1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-Ap3 and zinc finger protein ZFPap3"

ORIGIN
Query Match 97.8%; Score 3226.4; DB 6; Length 3300;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3254; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
Qy 61 GTCAAATTCAGGTGTGAATGTGAACACCGTAACGTTATACGATGTCCGAGATATCGCG 120
Db 61 GTCAAATTCAGGTGTGAATGTGAACACCGTAACGTTATACGATGTCCGAGATATCGCG 120
Qy 121 GTGTCTCTTATCAGACCGTTTCCCGGTGTGAACACCGAGCCAGCCAGTTCTGCGAAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGGTGTGTGAACACCGAGCCAGCCAGTTCTGCGAAAA 180
Qy 181 CCGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTTACATTTCCCAACCCGCTGGCAC 240
Db 181 CCGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTTACATTTCCCAACCCGCTGGCAC 240
Qy 241 AACAACTGCGCGGCAAAACAGTGTGCTGATTTGGCGGTGGCCACCTCCAGTGTGGCCCTGC 300
Db 241 AACAACTGCGCGGCAAAACAGTGTGCTGATTTGGCGGTGGCCACCTCCAGTGTGGCCCTGC 300
Qy 301 ACGCGCGGTGCGAAATGTGCGCGGCAATTAATCTCGCGCCGATCACTGCGGTGGCCAGCG 360
Db 301 ACGCGCGGTGCGAAATGTGCGCGGCAATTAATCTCGCGCCGATCACTGCGGTGGCCAGCG 360

Qy 361 TGGTGGTGTGCGATGCTAGAACGAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCACAATC 420
Db 361 TGGTGGTGTGCGATGCTAGAACGAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCACAATC 420
Qy 421 TTCTCGGCAACCGGTGAGTGGGCTGATCATTAATCTATCCGCTGGATGACCAAGATGCA 480
Db 421 TTCTCGGCAACCGGTGAGTGGGCTGATCATTAATCTATCCGCTGGATGACCAAGATGCA 480
Qy 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTCTGACCA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTCTGACCA 540
Qy 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 600
Qy 601 TGGTGCATTCGGGTGACCAAAATCGGCTGTTAGCGGGCCCAATTAAGTGTCTGTCTCGG 660
Db 601 TGGTGCATTCGGGTGACCAAAATCGGCTGTTAGCGGGCCCAATTAAGTGTCTGTCTCGG 660
Qy 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTTTCAACAAACCAATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTTTCAACAAACCAATGCAAAATGCTGA 780
Qy 781 ATGAGGCGATCGTTCCCACTCGCATGCTGGTGGCAACGATCAGATGCGCTGGCGCA 840
Db 781 ATGAGGCGATCGTTCCCACTCGCATGCTGGTGGCAACGATCAGATGCGCTGGCGCA 840
Qy 841 TGGCGCCCATACCGAGTCCGGGCTGGCGGTGGTGGCGATATCTCGGTAGTGGGATAG 900
Db 841 TGGCGCCCATACCGAGTCCGGGCTGGCGGTGGTGGCGATATCTCGGTAGTGGGATAG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAAACCACTCAACAGSATTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAAACCACTCAACAGSATTTC 960
Qy 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGTGTCGCACTCTCTCAGGCGCCAGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGTGTCGCACTCTCTCAGGCGCCAGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTGCCCGTCTCACTGTTGAAAAAAGAAAAACCACTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGCCCGTCTCACTGTTGAAAAAAGAAAAACCACTGGCGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGTGGCACGACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGTGGCACGACAGGTTT 1140
Qy 1141 CCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGAGTGTAGTCTCACTCATTTAG 1200
Db 1141 CCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGAGTGTAGTCTCACTCATTTAG 1200
Qy 1201 GCACAATTCATGTTTGACAGCTTATCATCGACTGCAAGCGGTCACCAATGCTCTGGCG 1260
Db 1201 GCACAATTCATGTTTGACAGCTTATCATCGACTGCAAGCGGTCACCAATGCTCTGGCG 1260
Qy 1261 TCAGGACCCATTCGGAAGCTGTGATGCGGTGTGAGTGTGAGTGTGTAATCACTGATTAATTCG 1320
Db 1261 TCAGGACCCATTCGGAAGCTGTGATGCGGTGTGAGTGTGAGTGTGTAATCACTGATTAATTCG 1320
Qy 1321 TGTGCTCAAGGCGCACTCCCGTTCTGGATTAATGTTTTTGGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGGCGCACTCCCGTTCTGGATTAATGTTTTTGGCGCGACATCAATACGGTT 1380
Qy 1381 CTGGCAAAATTTCTGAAATGAGTGTTCGCAATTAATCATCGGCTCGTATATGTTGTGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGTGTTCGCAATTAATCATCGGCTCGTATATGTTGTGA 1440
Qy 1441 ATTGTGAGCGGATACAAATTTTCACAGGAAACAGCCAGTCGCTTTAGGTGTTTTCACGA 1500

Db 1441 ATTGTGAGCGGATAACAATTTTACACAGAGAAACGCCAGTCGGTTAGGTGTTTTCACGA 1500
QY 1501 GCACCTTCCACCAACAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGTAATCTGG 1560
Db 1501 GCACCTTCCACCAACAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGTAATCTGG 1560
QY 1561 ATTAACCGCGATAAAGGCTCTAATACCGTCTCGCTGAAGTCGGTAAGAAAATTCGAGAAAGAT 1620
Db 1561 ATTAACCGCGATAAAGGCTCTAATACCGTCTCGCTGAAGTCGGTAAGAAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTACCGTCTGAGCATCCGATGATAAATCTGGAAGAGAAAATTCGAGGTT 1680
Db 1621 ACCGGAATTAAGTACCGTCTGAGCATCCGATGATAAATCTGGAAGAGAAAATTCGAGGTT 1680
QY 1681 GCGGCAATCGGCGATGGCCCTGACATTTCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAATCGGCGATGGCCCTGACATTTCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTTGGCTGAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCCGTACGTTTACAAACCGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCCGTACGTTTACAAACCGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
QY 1861 GAAGGTTATCGCTGATTATTAACAAGATCTGCTGCCGACAAAGCTGAAGAGCGGCTGATTTCAAC 1920
Db 1861 GAAGGTTATCGCTGATTATTAACAAGATCTGCTGCCGACAAAGCTGAAGAGCGGCTGATTTCAAC 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGTAAAGAGCGGCTGATTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGTAAAGAGCGGCTGATTTCAAC 1980
QY 1981 CTGCAAGAACGCTACTTACCTGGCGCTGATTGCTGCTGACGGGGTATTGCGTTCAAG 2040
Db 1981 CTGCAAGAACGCTACTTACCTGGCGCTGATTGCTGCTGACGGGGTATTGCGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGGCAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGGCAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACAACATGAATGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACAACATGAATGAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTAATAAGGCGAAGACGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAGGCGAAGACGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCAAATCGACACAGCAAAAGTGAATTTATGGTTAAACGGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACAGCAAAAGTGAATTTATGGTTAAACGGTACTGCCGACCTTC 2280
QY 2281 AAGGTCACCATCAACCGCTTGGTGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGTCACCATCAACCGCTTGGTGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2340
QY 2341 CCGAAACAAAGCTGGCAAAAGCTTCTCGAAAATCTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGCTGGCAAAAGCTTCTCGAAAATCTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGCTGAGTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGCTGAGTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGCGGAAAGATCCAGTATTGCGGCACACGATGAAAACCGCCGAGAAAGTGAATCATG 2520
Db 2461 TTGCGGAAAGATCCAGTATTGCGGCACACGATGAAAACCGCCGAGAAAGTGAATCATG 2520
QY 2521 CCGAAACATCCGCGAGATGTCGCTTTCTGTTATGTCGCTGGTATCTGCGGTGATCAAGCC 2580

Db 2521 CCGAAACATCCGCGAGATGTCGCTTTCTGGTATGTCGCTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAAATAAATAAACAACCTCGGGATCGAGGAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAAACAAATAAATAAACAACCTCGGGATCGAGGAAGGATTCAGAAATTC 2700
QY 2701 GGATCTCTTTCTGTCGGCCAGAGCGCCCTCGAGCCCGGGGAGAAAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTTCTGTCGGCCAGAGCGCCCTCGAGCCCGGGGAGAAAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTCGTAAGTCTTCTCTCAGAGCTCTCACTGCTGGCCACACAGCGTACCCAC 2820
Db 2761 CCGGAATGTCGTAAGTCTTCTCAGCAGAGCAGCTCCCTGGTGGGCCACACAGCGTACCCAC 2820
QY 2821 ACCGGTGAATAAACCGTATAAATGCCAGAGTGGGGCAAAATCTTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACCGGTGAATAAACCGTATAAATGCCAGAGTGGGGCAAAATCTTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGCTGCGCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCAGAAATGTGGC 2940
Db 2881 CTGCTGCGCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCAGAAATGTGGC 2940
QY 2941 AAGTCTTTCTCTCGGTCTGCAATCTCGTCCGGACCAACGTACTACACCGGGGAGAG 3000
Db 2941 AAGTCTTTCTGAGCAGTCCAGCAACCTGGTGGCCCAACCAAGTACTCACACCGGGGAGAG 3000
QY 3001 CCTATGCTTTGTCGGGAATGTTGTAAGTCTTTCAGCGCGAGCGATTAACCTGGTGGCCAC 3060
Db 3001 CCTATGCTTTGTCGGGAATGTTGTAAGTCTTTCAGCACCAAGTGGCTCTTTGGTTAGACAC 3060
QY 3061 CAGCGTACCCACACGCGGTGAAAAACCGTATAAATGCCAGAGTGGCAGAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACGCGGTGAAAAACCGTATAAATGCCAGAGTGGCAGAAATCTTTTAGC 3120
QY 3121 CAGCGCGCCACCTGCGGACCATCAACGCACTCATCTGCGGAGAGCCATCAAAATGT 3180
Db 3121 CAGCGCGCCACCTGCGGAGCCCATCAACGCACTCATCTGCGGAGAGCCATCAAAATGT 3180
QY 3181 CCAGAATGTGGCAAGTCTTTCTCTCGGTCTGCAAAATCTGTCGCGGACCAACCTGACTCAC 3240
Db 3181 CCAGAATGTGGCAAGTCTTTCTCAACTTCAGGCAACTTGGTCCGTCAACCAAGTACTCAC 3240
QY 3241 ACCGGTAAAAAATACTAGTGGCCAGCGCGGCGGAGTACCCGTACGAGTTCGCGACTACGT 3300
Db 3241 ACCGGTAAAAAATACTAGTGGCCAGCGCGGCGGAGTACCCGTACGAGTTCGCGACTACGT 3300

RESULT 4

AX202429 LOCUS AX202429 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 17 from Patent WO0152620.
ACCESSION AX202429
VERSION AX202429.1 GI:15392177

SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.

REFERENCE

1.
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 17 26-JUL-2001;
The Scripps Research Institute (US); SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)

FEATURES

Location/Qualifiers
source
1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-m4 and zinc finger protein

ORIGIN		2ZFPm4"	
Query Match		97.4%; Score 3215.2; DB 6; Length 3300;	
Best Local Similarity		98.4%; Pred. No. 0;	
Matches 3247; Conservative		0; Mismatches 53; Indels 0; Gaps 0;	
QY	1	CGGACACCATCAATGGTGGCAAAACCTTTCGGCGGTATGCGCATGATAGGCCCGGAAGAGA	60
DB	1	CGGACACCATCAATGGTGGCAAAACCTTTCGGCGGTATGCGCATGATAGGCCCGGAAGAGA	60
QY	61	GTCAATTACAGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGCAGAGTATGCG	120
DB	61	GTCAATTACAGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGCAGAGTATGCG	120
QY	121	GTGCTCTTATCAGACCGGTTTCCCGCGTGGTGAACACGCGCACGCTTTCTCGGAANA	180
DB	121	GTGCTCTTATCAGACCGGTTTCCCGCGTGGTGAACACGCGCACGCTTTCTCGGAANA	180
QY	181	CGCGGGAANAAGTGAAGCGCGATGGCGAGCTGAATTTACATTTCCCAACCCGCTGGCAC	240
DB	181	CGCGGGAANAAGTGAAGCGCGATGGCGAGCTGAATTTACATTTCCCAACCCGCTGGCAC	240
QY	241	AACAACTGCGCGGCAAAACAGTGTGCTGATTGGCGTTGSCACCTCCAGTCTGSCCCTGC	300
DB	241	AACAACTGCGCGGCAAAACAGTGTGCTGATTGGCGTTGSCACCTCCAGTCTGSCCCTGC	300
QY	301	ACGCGCGCTCGCAAAATGTCGCGGGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG	360
DB	301	ACGCGCGCTCGCAAAATGTCGCGGGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG	360
QY	361	TGGTGTGTCGATGTTAGAACGAAGCGGCTGAAAGCCGTGAAGCGGCGGTGCAATTC	420
DB	361	TGGTGTGTCGATGTTAGAACGAAGCGGCTGAAAGCCGTGAAGCGGCGGTGCAATTC	420
QY	421	TTCTCGCGCAACGCTGAGTGGGCTGATCATTAATCTCCGCTGGATGACACAGGATGCCA	480
DB	421	TTCTCGCGCAACGCTGAGTGGGCTGATCATTAATCTCCGCTGGATGACACAGGATGCCA	480
QY	481	TTGCTGTGGAAGCTGCTGCAATAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA	540
DB	481	TTGCTGTGGAAGCTGCTGCAATAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA	540
QY	541	CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGCACTGGCGGTGAGCATC	600
DB	541	CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGCGACTGGCGGTGAGCATC	600
QY	601	TGGTGTGCTGAGTACAGCAAAATCGGCTGTAGCGGCGCCATTAAGTCTGTCTCGG	660
DB	601	TGGTGTGCTGAGTACAGCAAAATCGGCTGTAGCGGCGCCATTAAGTCTGTCTCGG	660
QY	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAAACGGGAAGCGGACGTGAGTGCATCTCCGGTTTCAACAAACCATGCAAAATGCTGA	780
DB	721	CGGAAACGGGAAGCGGACGTGAGTGCATCTCCGGTTTCAACAAACCATGCAAAATGCTGA	780
QY	781	ATGAGGGCATCTTCCCATCGGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCA	840
DB	781	ATGAGGGCATCTTCCCATCGGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCA	840
QY	841	TGCGGCCATTACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG	900
DB	841	TGCGGCCATTACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG	900
QY	901	ACGATACCGAAGACGCTCATGTATATCCCGCGGTAAACCAATCAACAGGATTTTC	960
DB	901	ACGATACCGAAGACGCTCATGTATATCCCGCGGTAAACCAATCAACAGGATTTTC	960
QY	961	GCCTGCTGGGGCAACACGCTGGACCGCTTGTGCAACTCTCTCAGGCCAGGCGGTGA	1020
DB	961	GCCTGCTGGGGCAACACGCTGGACCGCTTGTGCAACTCTCTCAGGCCAGGCGGTGA	1020
QY	1021	AGGCAATCAGCTGTGTGCCCTCTCAGTGTGAAAGAAAAACACCTCTGGCGCCCAATA	1080
DB	1021	AGGCAATCAGCTGTGTGCCCTCTCAGTGTGAAAGAAAAACACCTCTGGCGCCCAATA	1080
QY	1081	CGCAAAACCGCTCTCCCGCGGTTGGCGGATCAATTAATGCAGCTGGCAGCAGAGGTTT	1140
DB	1081	CGCAAAACCGCTCTCCCGCGGTTGGCGGATCAATTAATGCAGCTGGCAGCAGAGGTTT	1140
QY	1141	CCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG	1200
DB	1141	CCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG	1200
QY	1201	GCACAAATCTCATGTTGACAGCTTATCATCGACTGACGCGTGCACCAATGCTTCGGCG	1260
DB	1201	GCACAAATCTCATGTTGACAGCTTATCATCGACTGACGCGTGCACCAATGCTTCGGCG	1260
QY	1261	TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTGCTAAATCACTGCATAAATTCG	1320
DB	1261	TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTGCTAAATCACTGCATAAATTCG	1320
QY	1321	TGTGCTCAAGCGCACCTCCCGTCTCGATTAATGTTTTTGGCGCGACATCATACCGTT	1380
DB	1321	TGTGCTCAAGCGCACCTCCCGTCTCGATTAATGTTTTTGGCGCGACATCATACCGTT	1380
QY	1381	CTGGCAAAATATTTCTGAAATGAGCTTTCACAGGAAACAGCAGTCCGTTTAGGTGTTTCACGA	1440
DB	1381	CTGGCAAAATATTTCTGAAATGAGCTTTCACAGGAAACAGCAGTCCGTTTAGGTGTTTCACGA	1440
QY	1441	ATTGTGAGCGGATCAAAATTTTACAGGAAACAGCAGTCCGTTTAGGTGTTTCACGA	1500
DB	1441	ATTGTGAGCGGATCAAAATTTTACAGGAAACAGCAGTCCGTTTAGGTGTTTCACGA	1500
QY	1501	GCACCTTCAACCAAGGACCATAGATTATGAAAGCTGAAAGAGTAAACTGGTAATCTCG	1560
DB	1501	GCACCTTCAACCAAGGACCATAGATTATGAAAGCTGAAAGAGTAAACTGGTAATCTCG	1560
QY	1561	ATTAAACCGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT	1620
DB	1561	ATTAAACCGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT	1680
DB	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT	1680
QY	1681	GCGGCAACTGGCGATGGCGCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC	1740
DB	1681	GCGGCAACTGGCGATGGCGCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCTGTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGTAT	1800
DB	1741	GCTCAATCTGGCTGTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGTAT	1800
QY	1801	CGGTTTACCTGGGATGCGGTACGTTTACACCGCAAGCTGATTGCTTACCCGATCGCTGT	1860
DB	1801	CGGTTTACCTGGGATGCGGTACGTTTACACCGCAAGCTGATTGCTTACCCGATCGCTGT	1860
QY	1861	GAAAGGTTATCGCTCATTTTAAACAAAGATCTGTCGCGAACCCGCCAAAAACCTGGGAA	1920
DB	1861	GAAAGGTTATCGCTCATTTTAAACAAAGATCTGTCGCGAACCCGCCAAAAACCTGGGAA	1920
QY	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGGTGAAGAGCGCGCTGATGTTCAAC	1980
DB	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGGTGAAGAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTACCTGGCGCTGATTGCTGTGACGGGGGTTATGCGTTCAAG	2040
DB	1981	CTGCAAGAACCGTACTTACCTGGCGCTGATTGCTGTGACGGGGGTTATGCGTTCAAG	2040
QY	2041	TATGAAAACGGCAAGTACGACATTAAGACGTTGGGCGTGGATTAACGCTGGCGGAAAGCG	2100
DB	2041	TATGAAAACGGCAAGTACGACATTAAGACGTTGGGCGTGGATTAACGCTGGCGGAAAGCG	2100

```
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCGCTTTAATAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCGCTTTAATAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
QY 2221 GCATGGTCAACATCGACACAGCAAGTGAATTATGGTTAACGGTACTCGCGACCTTC 2280
Db 2221 GCATGGTCCACATCGACACAGCAAGTGAATTATGGTTAACGGTACTCGCGACCTTC 2280
QY 2281 AAGGTCACCAACCCAGTTCGTTGGCGTGTGAGCGCAGGTATTAAGCGCCGCCAGT 2340
Db 2281 AAGGTCACCAACCCAGTTCGTTGGCGTGTGAGCGCAGGTATTAAGCGCCGCCAGT 2340
QY 2341 CCGACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGTCG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGTCG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
QY 2461 TTGGCGAAAGATCCACGATTGTCGCCACCATGGAAGAAACGCCCGAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTGTCGCCACCATGGAAGAAACGCCCGAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTCGCGTATCAAGGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTCGCGTATCAAGGCC 2580
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAAACATTAACATAACAAACAACTCGGGATCGAGGAAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAAACATTAACATAACAAACAACTCGGGATCGAGGAAAGGATTCAGAAATTC 2700
QY 2701 GGATCTCTCTCTGTGGCCAGCGCCCTCGAGCCGCGGAGGAGCCCTATGCTGT 2760
Db 2701 GGATCTCTCTCTGTGGCCAGCGCCCTCGAGCCGCGGAGGAGCCCTATGCTGT 2760
QY 2761 CCGGAATGTGGTAAGTCTCTCTCAGAGCTCTCACCTGTGCGCCACACGCGTACCAC 2820
Db 2761 CCGGAATGTGGTAAGTCTCTCAGAGCTCTCACCTGTGCGCCACACGCGTACCAC 2820
QY 2821 ACGGGTGAAACCGGTATAATGCGCAGAGTGGCGCAATCTTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGGTGAAACCGGTATAATGCGCAGAGTGGCGCAATCTTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGGTGGCCATCAACGCACTACTATGCGGAGAGCCATCAATCTCCAGATGTGGC 2940
Db 2881 CTGGTGGCCATCAACGCACTACTATGCGGAGAGCCATCAATCTCCAGATGTGGC 2940
QY 2941 AAGTCTTTCTCTCGGTCTGCAATCTCTGTCGGCACCACCAAGTACTCACACCGGGGAGAAG 3000
Db 2941 AAGTCTTTCTAGTGATGTGCGTATCTTTGCGAGGCCACCAAGTACTCACACCGGGGAGAAG 3000
QY 3001 CCCTATCTGTTCGGATGTGGTAAGTCTCTTACGCGCAGCGATACCTGGTGGCCAC 3060
Db 3001 CCCTATCTGTTCGGATGTGGTAAGTCTCTTCTCAGAGCTCTACCTGGTGGCCAC 3060
QY 3061 CAGCGTACCCACAGCGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACAGCGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAATCTTTTAGC 3120
QY 3121 CAGCCCGGCACCTGGCCAGCATCAACGCACTCATATCTGCGGAGAGGCCATCAAAATGT 3180
Db 3121 CCGAGCGATAACCTGGTGGCGCCATCAACGCACTCATATCTGCGGAGAGGCCATCAAAATGT 3180
```

```
QY 3181 CCAGATGTGGCAAGTCTTTCTCTCGGTCTGACAACTCTGTCGGCACCACGTAACCTAC 3240
Db 3181 CCAGATGTGGCAAGTCTTTCTCTCGGTCTGACAACTCTGTCGGCACCACGTAACCTAC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGGCGCAGTACCCTGACGAGTTCGCGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGGCGCAGTACCCTGACGAGTTCGCGACTACGCT 3300

RESULT 5
AX202428
LOCUS AX202428 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 16 from Patent WO0152620.
ACCESSION AX202428
VERSION AX202428.1 GI:15392176
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 other sequences; artificial sequences.
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 16 26-JUL-2001;
The Scripps Research Institute (US) ; SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES
source Location/Qualifiers
1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="Partial sequence of pMal-m3 and zinc finger protein ZFPm3"
```

ORIGIN

Query Match 97.2%; Score 3207.2; DB 6; Length 3300;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

```
QY 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATATAGCGCCCGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATATAGCGCCCGAAGAGA 60
QY 61 GTCAAATTCAGGCTGGTGAATGTGAAACCAAGTATACGATGTGCGAGATATGCCG 120
Db 61 GTCAAATTCAGGCTGGTGAATGTGAAACCAAGTATACGATGTGCGAGATATGCCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACAGGCGACCGCTTTCTGCGAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACAGGCGACCGCTTTCTGCGAAA 180
QY 181 CCGCGGAAAAAGTGGAAAGCGCGATGCGGAGCTGAATTTACATTCCCAACCGCGTGGC 240
Db 181 CCGCGGAAAAAGTGGAAAGCGCGATGCGGAGCTGAATTTACATTCCCAACCGCGTGGC 240
QY 241 AACAACTGGCGGCAACAGTGTCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAACAGTGTCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACSCGCGTCCGAAATGTTCGGCGCATTAATCTCGCGCGATCACTGGGTGGCAGCG 360
Db 301 ACSCGCGTCCGAAATGTTCGGCGCATTAATCTCGCGCGATCACTGGGTGGCAGCG 360
QY 361 TGGTGTGTGATGGTAGAACGAGCGGTGAAAGCTGTAAAGCGGCGGTGCACAATC 420
Db 361 TGGTGTGTGATGGTAGAACGAGCGGTGAAAGCTGTAAAGCGGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGGTGATGATTAATATATCCGTGGATGACAGGATGCCA 480
Db 421 TTCTCGCGCAACCGGTGATGATTAATATATCCGTGGATGACAGGATGCCA 480
QY 481 TTGCTGTGAAGCTGCTGACCTAAATGTTCCCGGCTTATTTCTTGATGTCTTGAC 540
Db 481 TTGCTGTGAAGCTGCTGACCTAAATGTTCCCGGCTTATTTCTTGATGTCTTGAC 540
```

Qy	541	CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600
Db	541	CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600
Qy	601	TGTCGCATTTGGGTACACAGCAATCGCGTGTGTAGCGGGCCCAATTAAGTTCTGTCTCGG	660
Db	601	TGTCGCATTTGGGTACACAGCAATCGCGTGTGTAGCGGGCCCAATTAAGTTCTGTCTCGG	660
Qy	661	CGCGTCTGCGTCTGGCTGGGCGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Db	661	CGCGTCTGCGTCTGGCTGGGCGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Qy	721	CGGAAACGGGAAGCGCATGAGTGCCTATGTCGGTTCCTCAACAAACATCGAAATGCTGA	780
Db	721	CGGAAACGGGAAGCGCATGAGTGCCTATGTCGGTTCCTCAACAAACATCGAAATGCTGA	780
Qy	781	ATGAGGCATCGTTCCCATCTCGCATGCTGTGTGCAACGATCAGATGGCGCTGGGGCCAA	840
Db	781	ATGAGGCATCGTTCCCATCTCGCATGCTGTGTGCAACGATCAGATGGCGCTGGGGCCAA	840
Qy	841	TGCGCGCAATTACGAGTCCGGGCTGCGGTGTGTGCGGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGCGCAATTACGAGTCCGGGCTGCGGTGTGTGCGGATATCTCGGTAGTGGGATACG	900
Qy	901	ACGNATCGGAAGACAGCTCATGTTATATCCGCCGTTTAAACCATCAAAACAGGATTTTC	960
Db	901	ACGNATCGGAAGACAGCTCATGTTATATCCGCCGTTTAAACCATCAAAACAGGATTTTC	960
Qy	961	GCCTGCTGGGGCAAAACAGCGTGACCGCTTGCTGCAACTCTCTCAGGGCCACGCGGTGA	1020
Db	961	GCCTGCTGGGGCAAAACAGCGTGACCGCTTGCTGCAACTCTCTCAGGGCCACGCGGTGA	1020
Qy	1021	AGGCAATCAGCTGTTGCCGCTCTCATGTGTGAAAGAAAACACCTCGCGCCCAATA	1080
Db	1021	AGGCAATCAGCTGTTGCCGCTCTCATGTGTGAAAGAAAACACCTCGCGCCCAATA	1080
Qy	1081	CGAAAACGGCTCTCCCGCGGCTTGCGCGATTCATTAAATGCAGCTGGCAGCAGGTTT	1140
Db	1081	CGAAAACGGCTCTCCCGCGGCTTGCGCGATTCATTAAATGCAGCTGGCAGCAGGTTT	1140
Qy	1141	CCCGACTGGAAGCGGCGAGTGACGCCAAACGCCAATTAATGTAGTTAGCTCACTCATTAG	1200
Db	1141	CCCGACTGGAAGCGGCGAGTGACGCCAAACGCCAATTAATGTAGTTAGCTCACTCATTAG	1200
Qy	1201	GCACAAATTCATGTTTGACAGCTTATCATCGACTGACCGGTGCACAAATGCTCTGGCG	1260
Db	1201	GCACAAATTCATGTTTGACAGCTTATCATCGACTGACCGGTGCACAAATGCTCTGGCG	1260
Qy	1261	TCAGCAGCCATCGGAAGCTGTGATGCGTGTGCGAGTGGTAATACACATGCAATAATCG	1320
Db	1261	TCAGCAGCCATCGGAAGCTGTGATGCGTGTGCGAGTGGTAATACACATGCAATAATCG	1320
Qy	1321	TGTCGCTCAAGCGGCATCTCCGTTTCTGGATAATGTTTTTTCGCGCGACATCATAAACGGTT	1380
Db	1321	TGTCGCTCAAGCGGCATCTCCGTTTCTGGATAATGTTTTTTCGCGCGACATCATAAACGGTT	1380
Qy	1381	CTGCGAAATATCTGAAATAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGGA	1440
Db	1381	CTGCGAAATATCTGAAATAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGGA	1440
Qy	1441	ATTGTGAGCGGATAACAAATTTTCACAGGAAACAGCCAGTCCGTTTAGGTGTTTTACGA	1500
Db	1441	ATTGTGAGCGGATAACAAATTTTCACAGGAAACAGCCAGTCCGTTTAGGTGTTTTACGA	1500
Qy	1501	GCACCTTCACCAACAGGACCATAGATTATGAAAACTGAAAGAGTAACCTGGTAATCTGG	1560
Db	1501	GCACCTTCACCAACAGGACCATAGATTATGAAAACTGAAAGAGTAACCTGGTAATCTGG	1560
Qy	1561	ATTAAACGGCGATAAAGGCTATAACCGTCTCGCTGAACTCGGTGAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAAACGGCGATAAAGGCTATAACCGTCTCGCTGAACTCGGTGAAGAAATTCGAGAAAGAT	1620

QY	1621	ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCTGGAAGAGAAATTTCCCAACAGGTT	1681
DB	1621	ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCTGGAAGAGAAATTTCCCAACAGGTT	1680
QY	1681	CGGCAACTCGCGATGCGCCCTGCATTTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
DB	1681	CGGCAACTCGCGATGCGCCCTGCATTTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTTCCAGGACAAAGCTGAT	1800
DB	1741	GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTTCCAGGACAAAGCTGAT	1800
QY	1801	CCGTTTACCTGGGATGCGGTACGTTTAAACGGGCAAGCTGATTTACCCCGATCGCTGTT	1860
DB	1801	CCGTTTACCTGGGATGCGGTACGTTTAAACGGGCAAGCTGATTTACCCCGATCGCTGTT	1860
QY	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA	1920
DB	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA	1920
QY	1921	GAGATCCCGCGCTGTAATAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTTCAAC	1980
DB	1921	GAGATCCCGCGCTGTAATAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGACGGGGTTATGCGTTCAAG	2040
DB	1981	CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGACGGGGTTATGCGTTCAAG	2040
QY	2041	TATGAAACCGCAGTACGACATTAAGACGTGGCGTGGATTAACGCTGGCGGGAAGCG	2100
DB	2041	TATGAAACCGCAGTACGACATTAAGACGTGGCGTGGATTAACGCTGGCGGGAAGCG	2100
QY	2101	GCTCTGACCTTCTCGTTTGACCTGATTAAAAAACAACATGATGACAGACACCGGATTAC	2160
DB	2101	GCTCTGACCTTCTCGTTTGACCTGATTAAAAAACAACATGATGACAGACACCGGATTAC	2160
QY	2161	TCCATCGAGAAGCTGCTTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG	2220
DB	2161	TCCATCGAGAAGCTGCTTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG	2220
QY	2221	GCATGGTCCACATCGACACCGCAGCAAGTGAATATGGTGTAAACGCTACGCGACCTTC	2280
DB	2221	GCATGGTCCACATCGACACCGCAGCAAGTGAATATGGTGTAAACGCTACGCGACCTTC	2280
QY	2281	AAGGGTCAACCATCCAAACCGTTTGGTGGTGTCTGAGCGCAGGTATTAACCGCCAGT	2340
DB	2281	AAGGGTCAACCATCCAAACCGTTTGGTGGTGTCTGAGCGCAGGTATTAACCGCCAGT	2340
QY	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGTGACTGATGAAGGTCGT	2400
DB	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGTGACTGATGAAGGTCGT	2400
QY	2401	GAAGCGGTTAATAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTACGAGGAAGAG	2460
DB	2401	GAAGCGGTTAATAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTACGAGGAAGAG	2460
QY	2461	TTGCGGAAAGATCCACGCTATTGCGCCACCATGGAAAAACGCCAGAAAGGTGAATCATG	2520
DB	2461	TTGCGGAAAGATCCACGCTATTGCGCCACCATGGAAAAACGCCAGAAAGGTGAATCATG	2520
QY	2521	CCGAACATCCGCGAGATGTCGCTTTTCTGGTATGCCGTGGCTACTGCGGTGATCAACGCC	2580
DB	2521	CCGAACATCCGCGAGATGTCGCTTTTCTGGTATGCCGTGGCTACTGCGGTGATCAACGCC	2580
QY	2581	GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGAAAGACGGCAGACTAATTCGAGTCG	2640
DB	2581	GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGAAAGACGGCAGACTAATTCGAGTCG	2640
QY	2641	AACAAACAAACAAATAACAACAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC	2700
DB	2641	AACAAACAAACAAATAACAACAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC	2700
QY	2701	GGATCCTCTTCTGTGGCCGAGCGCGGCCCTCGAGCCGGGGGAGAAGCCTATGCTGTG	2760

```
Db 2701 GGATCCTCTTCTCTGTGGCCAGCGGCGCTCGAGCCGGGAGAGCCCTATCTGT 2760
Qy 2761 CCGGAATGGTAAGTCTTCTCAGAGCTCTCAGAGTGGTGGCCAGGAGTACCCAC 2820
Db 2761 CCGGAATGGTAAGTCTTCTCAGAGTGGTGGCCAGGAGTACCCAC 2820
Qy 2821 ACGGTGAAAAACCGTATAAATGCGGAGTGGCGCAAAATCTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGTGAAAAACCGTATAAATGCGGAGTGGCGCAAAATCTTTAGCCAGGCGCTCC 2880
Qy 2881 CTGGTGGCCATCAAGCACTCATCTGCGGAGAGGCAATACAAATGTCAGAAATGTGGC 2940
Db 2881 CTGGTGGCCATCAAGCACTCATCTGCGGAGAGGCAATACAAATGTCAGAAATGTGGC 2940
Qy 2941 AAGTCTTCTCTCGGTCTGACAACTCTGTCGGGACCAACGCTACTCACCGGGGAGAG 3000
Db 2941 AAGTCTTCTCGGATGTTGTAAGTCTTTCAGCCAGAGCAGTCCCTGGTGGCCAG 3000
Qy 3001 CCTATGTTGTCGGAATGTTGTAAGTCTTTCAGCCAGAGCAGTCCCTGGTGGCCAG 3060
Db 3001 CCTATGTTGTCGGAATGTTGTAAGTCTTTCAGCCAGAGCAGTCCCTGGTGGCCAG 3060
Qy 3061 CAGGTATCCCAACGGGTGAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTAGC 3120
Db 3061 CAGGTATCCCAACGGGTGAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTAGT 3120
Qy 3121 CAGGCGGCGCCACCTGGCCAGGCATCAAGCGCACTCATCTGCGGAGAGGCAATACAAATGT 3180
Db 3121 GACTGCGCGACCTTGTCTGCGCCTCAAGCGCACTCATCTGCGGAGAGGCAATACAAATGT 3180
Qy 3181 CCAGATGTGGCAAGTCTTCTCGGTGTCGCAATCTCGTCCGCGACCAACGTAAGT 3240
Db 3181 CCAGATGTGGCAAGTCTTCTCGGTGTCGCAATCTCGTCCGCGACCAACGTAAGT 3240
Qy 3241 ACCGTAAAAAATAGTGGCGGCGGCGGCGGAGTACCGTACGAGTTCGGAGTACGCT 3300
Db 3241 ACCGTAAAAAATAGTGGCGGCGGCGGCGGAGTACCGTACGAGTTCGGAGTACGCT 3300
```

```
RESULT 6
LOCUS AX284144
DEFINITION Sequence 1 from Patent WO0178756.
ACCESSION AX284144
VERSION AX284144.1 GI:17044832
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS Wiedersanders, B. and Maubach, G.
TITLE Agent for postoperative use after the removal of bone tumours
JOURNAL Patent: WO 0178756-A 1 25-OCT-2001;
Depuy Biotech Jena GmbH (DE)
FEATURES
source
1. .7373
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="Expressionsvektor pMalc2cchbnp2"
1528. 3453
/notes="unnamed protein product; /produkt=Fusionsprotein
MBP Cystatin C BMP-2"
/codon_start=1
/transl_table=11
/protein_id="CAD12160.1"
/db_xref="GI:17044833"
/translation="MKIEEKLIWINGDKYNGLAIEVKGKFEKDTGKIVTVHPDKL
EEKFPQVATGGDPIIFWAHDFRFGYAQGLLAEITPDKAFODKLYPFTVADRYNG
KLIAPIFAEALSLLIYNKDLPPPKTWEEIPALDKELAKSKSALMFLNLPYFTWP
LIAADGGYAFKNGYKIDVGVNAGAKAGLTFVLDTKNKHMNADTYSIAEAAF
NKGETAMTINGPWAWSNIDTSKVNYGVTVLPTFKGQPSKFPFVGLSAGINAASPNKEL
```

```
misc_feature
/notes="Faktor Xa Protease /position=388 (AMINO
ACIDsequenz)"
3067. .3108
/misc_feature
/notes="/spacer zwischen Cystatin C und BMP-2"
ORIGIN
Query Match 82.0%; Score 2705.8; DB 6; Length 7373;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2719; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGGCATAGAGCCCGGAGAGA 60
Db 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGGCATAGAGCCCGGAGAGA 60
Qy 61 GTCAATTCAGGGTGGTGAATGTGA AACCAAGTATACGATGTCCAGAGATATCCG 120
Db 61 GTCAATTCAGGGTGGTGAATGTGA AACCAAGTATACGATGTCCAGAGATATCCG 120
Qy 121 GTGCTCTTTATCAGACCGTTTCCGGGTGTGAACAGGCGCAGCCAGCTTTCTCGGAAAA 180
Db 121 GTGCTCTTTATCAGACCGTTTCCGGGTGTGAACAGGCGCAGCCAGCTTTCTCGGAAAA 180
Qy 181 CGCGGAAAAAAGTGAAGCGCGCATGCGGAGCTGAATACATTCCTCAACCGGCTGGCAC 240
Db 181 CGCGGAAAAAAGTGAAGCGCGCATGCGGAGCTGAATACATTCCTCAACCGGCTGGCAC 240
Qy 241 AACAACTGGCGGCAACAGTCTGATTCGGCGTTGCGACCTCCAGTCTCGCCCTGC 300
Db 241 AACAACTGGCGGCAACAGTCTGATTCGGCGTTGCGACCTCCAGTCTCGCCCTGC 300
Qy 301 ACGCGCGCTCGCAAAATGTTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCAGCG 360
Db 301 ACGCGCGCTCGCAAAATGTTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCAGCG 360
Qy 361 TGGTGGTGTGATGGTAGAACGAGCGGCTCGAAGCGCTGTAAGCGGCTGCACAATC 420
Db 361 TGGTGGTGTGATGGTAGAACGAGCGGCTCGAAGCGCTGTAAGCGGCTGCACAATC 420
Qy 421 TTCTCGCGCAACCGCTCAGTGGCTGATTAATCACTATCCGCTGGATGACAGAGTCCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGCTGATTAATCACTATCCGCTGGATGACAGAGTCCCA 480
Qy 481 TTGCTGTGGAAGCTGCTGCACTAATTTTCGGCGGTTATTTCTTGATGCTCTGACGAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAATTTTCGGCGGTTATTTCTTGATGCTCTGACGAGA 540
Qy 541 CACCCATCAACAGTATTTATTTTCTCCATGAAGAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTATTTTCTCCATGAAGAGCGGTACGCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTTCGCAATGGGTACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTCTGTCTCGG 660
Db 601 TGGTTCGCAATGGGTACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTCTGTCTCGG 660
Qy 661 CCGGTCTCGGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCCATAG 720
Db 661 CCGGTCTCGGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCCATAG 720
Qy 721 CGGAACGGGAAGCGGACTGGAGTGCAGTCCGGTTTCAACAAACCATCTGCAATCTCTGA 780
Db 721 CGGAACGGGAAGCGGACTGGAGTGCAGTCCGGTTTCAACAAACCATCTGCAATCTCTGA 780
Qy 781 ATGAGGGCATCGTTCCCACTCGGTGTTGCCAAGCATCAGATGGCGCTGGGGCGAA 840
Db 781 ATGAGGGCATCGTTCCCACTCGGTGTTGCCAAGCATCAGATGGCGCTGGGGCGAA 840
```

Qy	841	TGCGGCCATTACCGAGTCCGGGCTGCGCTGGTGGGATATCTCGGTAGTGGGATACG	900
Db			
Db	841	TGCGGCCATTACCGAGTCCGGGCTGCGCTGGTGGGATATCTCGGTAGTGGGATACG	900
Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCGCCGTTAAACCAACATCAACAGATTTTC	960
Db			
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCGCCGTTAAACCAACATCAACAGATTTTC	960
Qy	961	GCCTGCTGGGGCAAAACAGCGGTGACCCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
Db			
Db	961	GCCTGCTGGGGCAAAACAGCGGTGACCCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
Qy	1021	AGGCAATCAGCTGTTGCCGCTCTCATCTGGTGAAGAAAAACAACCTCGCGCCCAATYA	1080
Db			
Db	1021	AGGCAATCAGCTGTTGCCGCTCTCATCTGGTGAAGAAAAACAACCTCGCGCCCAATYA	1080
Qy	1081	CGCAAAACCGGCTCTCCCGCGGCTTGCGCGATCATTAATGCAGCTGGCAGCACAGGTTT	1140
Db			
Db	1081	CGCAAAACCGGCTCTCCCGCGGCTTGCGCGATCATTAATGCAGCTGGCAGCACAGGTTT	1140
Qy	1141	CCCAGCTGGAAAGCGGCGAGTGAGCGCAACGCCAATTAATGTGAGTTAGCTCACTATTAG	1200
Db			
Db	1141	CCCAGCTGGAAAGCGGCGAGTGAGCGCAACGCCAATTAATGTGAGTTAGCTCACTATTAG	1200
Qy	1201	GCACAAATCTCATGTTTTCAGAGCTTATCATGAGCTGCACGGTGACCAATGCTTTGGCG	1260
Db			
Db	1201	GCACAAATCTCATGTTTTCAGAGCTTATCATGAGCTGCACGGTGACCAATGCTTTGGCG	1260
Qy	1261	TCAGGACGCCAATCGGAAGCTGTGTTATGGCTGTGCAGTCTGTAATCACATGCATAATTCG	1320
Db			
Db	1261	TCAGGACGCCAATCGGAAGCTGTGTTATGGCTGTGCAGTCTGTAATCACATGCATAATTCG	1320
Qy	1321	TGTCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCATAACGGTT	1380
Db			
Db	1321	TGTCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCATAACGGTT	1380
Qy	1381	CTGGCAAAATTTCTGAAATGAGCTGTTTGACAAATTAATCATCGGCTCGTATAATGTGGGA	1440
Db			
Db	1381	CTGGCAAAATTTCTGAAATGAGCTGTTTGACAAATTAATCATCGGCTCGTATAATGTGGGA	1440
Qy	1441	ATTGTGAGCGGATAACAAATTTACACAGGAAACAGCCAGTCCGTTTAGTGTTTTTCACGA	1500
Db			
Db	1441	ATTGTGAGCGGATAACAAATTTACACAGGAAACAGCCAGTCCGTTTAGTGTTTTTCACGA	1500
Qy	1501	GCATTTACCAACAAAGGACCATAGATTATGAAAACTGAAGAAGTAAACTGGTAATCTGG	1560
Db			
Db	1501	GCATTTACCAACAAAGGACCATAGATTATGAAAACTGAAGAAGTAAACTGGTAATCTGG	1560
Qy	1561	ATTTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAGAATTTTCGAGAAAGAT	1620
Db			
Db	1561	ATTTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAGAATTTTCGAGAAAGAT	1620
Qy	1621	ACCGGAATTTAAGTTCACGGTTGAGCATCCGGATAAACTGGAAGAGAAATTTCCACAGGTT	1680
Db			
Db	1621	ACCGGAATTTAAGTTCACGGTTGAGCATCCGGATAAACTGGAAGAGAAATTTCCACAGGTT	1680
Qy	1681	CGCGCAACTGCGCATGCGCCCTGCAATTAATCTTCTGGGCAACGACCGTTTGGTGGCTAC	1740
Db			
Db	1681	CGCGCAACTGCGCATGCGCCCTGCAATTAATCTTCTGGGCAACGACCGTTTGGTGGCTAC	1740
Qy	1741	GCTCAATCTGGCCTGTTGGCTGGAATCACCCCGGACAAAGGTTTCCAGGACAGCTGTAT	1800
Db			
Db	1741	GCTCAATCTGGCCTGTTGGCTGGAATCACCCCGGACAAAGGTTTCCAGGACAGCTGTAT	1800
Qy	1801	CCGTTTACCTGGGATGCGGTACGTTTAAACCGGCAAGCTGATTGCTTACC CGCATCGCTGTT	1860
Db			
Db	1801	CCGTTTACCTGGGATGCGGTACGTTTAAACCGGCAAGCTGATTGCTTACC CGCATCGCTGTT	1860
Qy	1861	GAAGCGTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAACCCCGCAAAAACCTGGGAA	1920
Db			
Db	1861	GAAGCGTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAACCCCGCAAAAACCTGGGAA	1920

Qy	1921	GAGATCCGGCGTGGATAAAGAACTGAAAGCGAAGAGGTAAAGAGCGCGCTGATGTTTCAAC	1980
Db	1921	GAGATCCGGCGCTGGATAAAGAACTGAAAGCGAAGAGGTAAAGAGCGCGCTGATGTTTCAAC	1980
Qy	1981	CTGCAAGAACCGTACTTTCACCTGSCCGCTGATTGCTCTGACGGGGGTATGCGTTCAG	2040
Db	1981	CTGCAAGAACCGTACTTTCACCTGSCCGCTGATTGCTCTGACGGGGGTATGCGTTCAG	2040
Qy	2041	TATGAAAAACGCAAGTAGTACGACATTAAAGACGTGGCGCTGGATAAACGTGCGCGAAGACG	2100
Db	2041	TATGAAAAACGCAAGTAGTACGACATTAAAGACGTGGCGCTGGATAAACGTGCGCGAAGACG	2100
Qy	2101	GGTCTGACCTTCGTGGTGAACCTGATTAAAAAANAACACATGATGACGACACCGATTAC	2160
Db	2101	GGTCTGACCTTCGTGGTGAACCTGATTAAAAAANAACACATGATGACGACACCGATTAC	2160
Qy	2161	TCCATCGAAGAGCTGCGCTTTAATAAAGGCGAAACACGCGATGACCATCAACGGCGCCGTGG	2220
Db	2161	TCCATCGAAGAGCTGCGCTTTAATAAAGGCGAAACACGCGATGACCATCAACGGCGCCGTGG	2220
Qy	2221	GCATGGTCCAACTCGACACAGCAAAAGTGAATTTATGGTGAACGGTACTGCGCACCTTC	2280
Db	2221	GCATGGTCCAACTCGACACAGCAAAAGTGAATTTATGGTGAACGGTACTGCGCACCTTC	2280
Qy	2281	AAGGGTCAAACCATCCAAACCGTTCGTTGGCGGTGCTGAGCGCAGGTATTAACGCCGCCAGT	2340
Db	2281	AAGGGTCAAACCATCCAAACCGTTCGTTGGCGGTGCTGAGCGCAGGTATTAACGCCGCCAGT	2340
Qy	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGCTCG	2400
Db	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGCTCG	2400
Qy	2401	GAAAGCGGTTAATAAGACAAACCCCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG	2460
Db	2401	GAAAGCGGTTAATAAGACAAACCCCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG	2460
Qy	2461	TTGGCGAAAGATCCACGTATTGCGCGCACCATGAAACCGCCAGAAAGGTGAAATCATG	2520
Db	2461	TTGGCGAAAGATCCACGTATTGCGCGCACCATGAAACCGCCAGAAAGGTGAAATCATG	2520
Qy	2521	CCGAACATCCCGCAGATGTCCGCTTTCTGGTATCCCGTGCGTACTCGGGTGATCAAGCC	2580
Db	2521	CCGAACATCCCGCAGATGTCCGCTTTCTGGTATCCCGTGCGTACTCGGGTGATCAAGCC	2580
Qy	2581	GCCAGCGGTGCTGACTGTGATGAAGCCCTGAAAGACGGCAGACTTAATTCGAGCTCG	2640
Db	2581	GCCAGCGGTGCTGACTGTGATGAAGCCCTGAAAGACGGCAGACTTAATTCGAGCTCG	2640
Qy	2641	AACAACAACAACATTAACAATAACAACCTCTGGGATCGAGGGAAGGATTTTCAGAAATTC	2700
Db	2641	AACAACAACAACATTAACAATAACAACCTCTGGGATCGAGGGAAGGATTTTCAGAAATTC	2700
Qy	2701	GGATCCTCTTCTCTGTGGCCACAGCGCGCTTCGAGCCCGG	2741
Db	2701	GGATCCTCTCTCTCGGCAAGCCGCGCTCTGGTGGAGG	2741

[illegible]

FEATURES		Location/Qualifiers	ORIGIN
source			
(FR)			
1. .8101			
/organism="synthetic construct"			
/mol_type="unassigned DNA"			
/db_xref="taxon:32630"			
/note="bouche JM109 / pMAL - Hs Cdc25C"			
Query Match		82.0%; Score 2705.8; DB 6; Length 8101;	
Best Local Similarity		99.9%; Pred. No. 0;	
Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;			
QY	1	CGACACCATCGAATGGTGTCAAAACCTTTTCGGGTATGGCATATAGCGCCCGAAGAGA	60
Db	1	CGACACCATCGAATGGTGTCAAAACCTTTTCGGGTATGGCATATAGCGCCCGAAGAGA	60
QY	61	GTCAATTCAGGGTGTGAATGTGAACCAAGTAACGTTATACGATGTCCAGAGTATGCCG	120
Db	61	GTCAATTCAGGGTGTGAATGTGAACCAAGTAACGTTATACGATGTCCAGAGTATGCCG	120
QY	121	GTGTCCTTATCAGACCGCTTTCCCGGTGGTGAACCGCCAGCCAGCTTTCTCGRAAA	180
Db	121	GTGTCCTTATCAGACCGCTTTCCCGGTGGTGAACCGCCAGCCAGCTTTCTCGRAAA	180
QY	181	CGCGGAAAAAGTGAACCGCGATGGCGAGCTGAATTAACATTCCTCCAAACCGCTGGCAC	240
Db	181	CGCGGAAAAAGTGAACCGCGATGGCGAGCTGAATTAACATTCCTCCAAACCGCTGGCAC	240
QY	241	AACAATCGCGGCAACAGTCGTGCTGATGGGTTGCCACTCCAGTTCGCCCTGC	300
Db	241	AACAATCGCGGCAACAGTCGTGCTGATGGGTTGCCACTCCAGTTCGCCCTGC	300
QY	301	ACGCGCGTGCACAAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG	360
Db	301	ACGCGCGTGCACAAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG	360
QY	361	TGTTGGTGTGATGTAGAACGAAGCGCGTGCAGAGCTGTAAAGCGCGGTGCACAATC	420
Db	361	TGTTGGTGTGATGTAGAACGAAGCGCGTGCAGAGCTGTAAAGCGCGGTGCACAATC	420
QY	421	TTCTCGGCAACCGGTACAGTCGGGTGATCAATTAATCTCGCGCGGATCAACTGGGTGCCAGTCCCA	480
Db	421	TTCTCGGCAACCGGTACAGTCGGGTGATCAATTAATCTCGCGCGGATCAACTGGGTGCCAGTCCCA	480
QY	481	TTGCTGTGGAAGTCCCTGCACTAATTTTCGGCGGTTAATTTCTTGATGTCTCTGACCCAGA	540
Db	481	TTGCTGTGGAAGTCCCTGCACTAATTTTCGGCGGTTAATTTCTTGATGTCTCTGACCCAGA	540
QY	541	CACCATCAACAGTATTAATTTTCCCATGAAGACGGTACGGCTAGCGGTGGAGCATC	600
Db	541	CACCATCAACAGTATTAATTTTCCCATGAAGACGGTACGGCTAGCGGTGGAGCATC	600
QY	601	TGTTGCGAATTTGGGTACACAGCAAAATCGCTGTAGCGGGCCCAATTAAGTTCTGTCGCG	660
Db	601	TGTTGCGAATTTGGGTACACAGCAAAATCGCTGTAGCGGGCCCAATTAAGTTCTGTCGCG	660
QY	661	CGGCTCTGCGTCTGGCTGGCTGGAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Db	661	CGGCTCTGCGTCTGGCTGGCTGGAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAACGGGAAGGCACTGGAGTCCCATGTCGGTTTTTCAACAAACCATGCAATGCTGA	780
Db	721	CGGAACGGGAAGGCACTGGAGTCCCATGTCGGTTTTTCAACAAACCATGCAATGCTGA	780
QY	781	ATGAGGCACTGTTTCCCACTCGCATGCTGTTGGCAACAGTCAAGTGGCGCTGGGCGCAA	840
Db	781	ATGAGGCACTGTTTCCCACTCGCATGCTGTTGGCAACAGTCAAGTGGCGCTGGGCGCAA	840
QY	841	TGCGCGCATTAACGAGTCCGGGTGCGGTGGTGGGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGCGCATTAACGAGTCCGGGTGCGGTGGTGGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACACCATCAAAACAGATTTTC	960
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACACCATCAAAACAGATTTTC	960
QY	961	GCCTGCTGGGGCAAAACAGCGGTGGACCGCTTCTCAGGGGCGAGCGGTGA	1020
Db	961	GCCTGCTGGGGCAAAACAGCGGTGGACCGCTTCTCAGGGGCGAGCGGTGA	1020
QY	1021	AGGCAATCAGCTGTGTCCTCTCACTGTGTAAGAAAAACACCTCGGCGCCCAATA	1080
Db	1021	AGGCAATCAGCTGTGTCCTCTCACTGTGTAAGAAAAACACCTCGGCGCCCAATA	1080
QY	1081	CGAAAAACCGCTCTCCCGCGCTTGGCCGATTCATTAATGCGAGCTGGCAGCAGAGTTT	1140
Db	1081	CGAAAAACCGCTCTCCCGCGCTTGGCCGATTCATTAATGCGAGCTGGCAGCAGAGTTT	1140
QY	1141	CCGCACTGGAAGCGGCGAGTGAACGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG	1200
Db	1141	CCGCACTGGAAGCGGCGAGTGAACGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG	1200
QY	1201	GCACATTCATGTTTTCAGAGCTTATCATCGAGTGCACGGTGCACCAATGCTTCTGGCG	1260
Db	1201	GCACATTCATGTTTTCAGAGCTTATCATCGAGTGCACGGTGCACCAATGCTTCTGGCG	1260
QY	1261	TCAGGACAGCCATCGGAAGCTGTGATGGCTGTGCAAGTGTAAATCACATGCATAATTCG	1320
Db	1261	TCAGGACAGCCATCGGAAGCTGTGATGGCTGTGCAAGTGTAAATCACATGCATAATTCG	1320
QY	1321	TGTCGCTCAAGGGCGACTCCCGTTCTGGATAATGTTTTTTCGCCCGACATCAACCGGTT	1380
Db	1321	TGTCGCTCAAGGGCGACTCCCGTTCTGGATAATGTTTTTTCGCCCGACATCAACCGGTT	1380
QY	1381	CTGSCAAATTCCTGAAATGAGCTGTGCAATTAATCATCGGCTGGTAAATGTGTGGA	1440
Db	1381	CTGSCAAATTCCTGAAATGAGCTGTGCAATTAATCATCGGCTGGTAAATGTGTGGA	1440
QY	1441	ATTGTGAGCGGATAACAAATTTTCAACAGGAAACAGCAGTCCGTTTTCAGTGTTCACGA	1500
Db	1441	ATTGTGAGCGGATAACAAATTTTCAACAGGAAACAGCAGTCCGTTTTCAGTGTTCACGA	1500
QY	1501	GCATTTCAACCAACAGGACCATAGATATGAAATCTGAAAGAGGTAAATCTGGTAAATCTCG	1560
Db	1501	GCATTTCAACCAACAGGACCATAGATATGAAATCTGAAAGAGGTAAATCTGGTAAATCTCG	1560
QY	1561	ATTAAACGGGATAAAGCTTAAACGGTCTCGCTGAGTAAAGAAATTCGGAAGAT	1620
Db	1561	ATTAAACGGGATAAAGCTTAAACGGTCTCGCTGAGTAAAGAAATTCGGAAGAT	1620
QY	1621	ACCGGAATTAAGTCAACCGTTCAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGGTT	1680
Db	1621	ACCGGAATTAAGTCAACCGTTCAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGGTT	1680
QY	1681	GGGCAACTGGGATGGCCCTTGACATTTCTTGGGCAACAGCCGCTTTGGTGGCTAC	1740
Db	1681	GGGCAACTGGGATGGCCCTTGACATTTCTTGGGCAACAGCCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCTGTGTGCTGAAATCACCCGCAACAAAGCGTTCCAGGACAAAGCTGAT	1800
Db	1741	GCTCAATCTGGCTGTGTGCTGAAATCACCCGCAACAAAGCGTTCCAGGACAAAGCTGAT	1800
QY	1801	CGTTTTACTTGGATGCCGTACGTTACAAACGCAAGCTGATTCGTTACCGGATCGCTGTT	1860
Db	1801	CGTTTTACTTGGATGCCGTACGTTACAAACGCAAGCTGATTCGTTACCGGATCGCTGTT	1860
QY	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGTCGCAACCCGCAAAACCTGGGAA	1920
Db	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGTCGCAACCCGCAAAACCTGGGAA	1920
QY	1921	GAGATCCCGCGCTGGATAAGAACTGAAAGCGGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATAAGAACTGAAAGCGGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTACCTGGCGCTGATGCTGTGACGCGGGGTTATGCGTTTCAAG	2040

Db	1981	CTGCAAGAACCGTACTTTCACCTGGCGGTGATGCTGTGACGGGGTATGCGTTCAG	2040	
Qy	2041	TATGAAAAACGGCAAGTACGACATTAAAGACGTGGCGGTGGATAACCGCTGGCGGAAAGCG	2100	
Db	2041	TATGAAAAACGGCAAGTACGACATTAAAGACGTGGCGGTGGATAACCGCTGGCGGAAAGCG	2100	
Qy	2101	GGTCTGACCTTCTCTGGTGAACCTGATTAATAAACAACACATGAATGCGAGACACCGATTAC	2160	
Db	2101	GGTCTGACCTTCTCTGGTGAACCTGATTAATAAACAACACATGAATGCGAGACACCGATTAC	2160	
Qy	2161	TCCATCGCAGAGCTGCCCTTTAATAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG	2220	
Db	2161	TCCATCGCAGAGCTGCCCTTTAATAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG	2220	
Qy	2221	GCATGGTCCAAATCGACACACGCAAGAGTGAATTAATGGTGTAAACGGTACTGCGGACCTTC	2280	
Db	2221	GCATGGTCCAAATCGACACACGCAAGAGTGAATTAATGGTGTAAACGGTACTGCGGACCTTC	2280	
Qy	2281	AAGGTCAACCATCAAAACCGTTCGTGGCGTGTGAGCGAGGTATTAACCGCCCGCAGT	2340	
Db	2281	AAGGTCAACCATCAAAACCGTTCGTGGCGTGTGAGCGAGGTATTAACCGCCCGCAGT	2340	
Qy	2341	CGAACAAAGAGCTGGCAAAAGAGTTCTCGAARACTATCTGCTGACTGATGAAGGTCTG	2400	
Db	2341	CGAACAAAGAGCTGGCAAAAGAGTTCTCGAARACTATCTGCTGACTGATGAAGGTCTG	2400	
Qy	2401	GAAGCGGTTAATAAGACACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAG	2460	
Db	2401	GAAGCGGTTAATAAGACACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAG	2460	
Qy	2461	TTGGCGAAAGATCCACGATATTCGCCGCCACCATGGAAGAACGCCAGAAAGGTGAATCATG	2520	
Db	2461	TTGGCGAAAGATCCACGATATTCGCCGCCACCATGGAAGAACGCCAGAAAGGTGAATCATG	2520	
Qy	2521	CGAACATCCCGCAGATGTCGCTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCC	2580	
Db	2521	CGAACATCCCGCAGATGTCGCTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCC	2580	
Qy	2581	GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGAAAGACGCCGACACTAAATTCGAGCTCG	2640	
Db	2581	GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGAAAGACGCCGACACTAAATTCGAGCTCG	2640	
Qy	2641	AACAACAACAATAACAATAACAACACTCGGGATCGAGGGAAGATTTCAGAAATTC	2700	
Db	2641	AACAACAACAATAACAATAACAACACTCGGGATCGAGGGAAGATTTCAGAAATTC	2700	
Qy	2701	GGATCCTCT	2709	
Db	2701	GGATCCTCT	2709	
RESULT 8				
AX377531				
LOCUS				
DEFINITION				
Sequence 8 from Patent WO0212553.				
ACCESSION				
AX377531				
VERSION				
AX377531.1 GI:19573717				
KEYWORDS				
synthetic construct				
synthetic construct				
other sequences; artificial sequences.				
SOURCE				
ORGANISM				
1				
REFERENCE				
AUTHORS				
Kappel,A., Polakowski,T., Pignot,M., Windhab,N., Behrensorf,H. and				
Muth,J.				
METHOD				
for detecting mutations in nucleotide sequences				
TITLE				
Patent: WO 0212553-A 8 14-FEB-2002;				
JOURNAL				
Nanogen Recognomics GmbH (DE)				
FEATURES				
Location/Qualifiers				
1..6648				
/organism="synthetic construct"				
/mol_type="unassigned DNA"				
/db_xref="taxon:32630"				
source				

/note="Plasmid pMAL-c2x"				
ORIGIN				
Query Match				
Best Local Similarity 81.8%; Score 2701; DB 6; Length 6648;				
Matches 2704; Conservative 0; Mismatches 5; Indels 0; Gaps 0;				
Qy	1	CCGACACCATCGAATGGTGTGCAAAACCTTTTCGGGTATGCGCATGATAGCCCGGAAGAGA	60	
Db	1	CCGACACCATCGAATGGTGTGCAAAACCTTTTCGGGTATGCGCATGATAGCCCGGAAGAGA	60	
Qy	61	GTCAATTTCAGGGTGTGAATGTGAACCCAGTAAAGTATATAGATGTCCAGAGTATCGG	120	
Db	61	GTCAATTTCAGGGTGTGAATGTGAACCCAGTAAAGTATATAGATGTCCAGAGTATCGG	120	
Qy	121	GTGTCTCTTATCAGACCGTTCCTCGCGGTGTGAACCCAGCCAGCCACGCTTTCTCGAAAA	180	
Db	121	GTGTCTCTTATCAGACCGTTCCTCGCGGTGTGAACCCAGCCAGCCACGCTTTCTCGAAAA	180	
Qy	181	CGCGGAAAAAGTGGAGCGCGATGCGGAGCTGAATTAATTCCTCCAAACCGCTGGCAC	240	
Db	181	CGCGGAAAAAGTGGAGCGCGATGCGGAGCTGAATTAATTCCTCCAAACCGCTGGCAC	240	
Qy	241	AACAACTGGCGGCAAAACAGTCTGTGATTTGGGCTTGCACCTCCAGTCTGGCCCTGC	300	
Db	241	AACAACTGGCGGCAAAACAGTCTGTGATTTGGGCTTGCACCTCCAGTCTGGCCCTGC	300	
Qy	301	ACGCGCCCTCGCAAAATTTGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCGCAGG	360	
Db	301	ACGCGCCCTCGCAAAATTTGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCGCAGG	360	
Qy	361	TGGTGGTGTGATGGTAGAAAGAGCGCGTTCGAAGCTGTAAAGCGCGGTGCAAAATC	420	
Db	361	TGGTGGTGTGATGGTAGAAAGAGCGCGTTCGAAGCTGTAAAGCGCGGTGCAAAATC	420	
Qy	421	TTCTCGCGCAACGGTCTAGTGGCTGATCAATTAATCTCCGTGATGATGATGATGATGATG	480	
Db	421	TTCTCGCGCAACGGTCTAGTGGCTGATCAATTAATCTCCGTGATGATGATGATGATGATG	480	
Qy	481	TTGTCTGTGGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	540	
Db	481	TTGTCTGTGGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	540	
Qy	541	CACCCATCAACAGTATTAATTTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600	
Db	541	CACCCATCAACAGTATTAATTTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600	
Qy	601	TGGTGGCATTTGGTCCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCCTGCTCGG	660	
Db	601	TGGTGGCATTTGGTCCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCCTGCTCGG	660	
Qy	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720	
Db	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720	
Qy	721	CGGAAACGGGAAGCGACTGGAGTGCATGTCCGTTTTTCAACAAACCATGCAAAATGCTGA	780	
Db	721	CGGAAACGGGAAGCGACTGGAGTGCATGTCCGTTTTTCAACAAACCATGCAAAATGCTGA	780	
Qy	781	ATGAGGGCATGTTTCCCACTCGGATGCTGTTGTCGCAACGATGATGATGATGATGATGATG	840	
Db	781	ATGAGGGCATGTTTCCCACTCGGATGCTGTTGTCGCAACGATGATGATGATGATGATGATG	840	
Qy	841	TGCGGCCATTTACCGAGTCCGGCTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTG	900	
Db	841	TGCGGCCATTTACCGAGTCCGGCTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTG	900	
Qy	901	ACGATACCGGAAGACAGCTCATGTTTATATCCCGCGGTGATTAACCAACCATCAAAACGAGT	960	
Db	901	ACGATACCGGAAGACAGCTCATGTTTATATCCCGCGGTGATTAACCAACCATCAAAACGAGT	960	
Qy	961	GCCTGCTGGGGCAACCAAGCGGTGAGACCGCTTGTGCTGCAACTCTCTCAGGCGCAGGCGGTGA	1020	
Db	961	GCCTGCTGGGGCAACCAAGCGGTGAGACCGCTTGTGCTGCAACTCTCTCAGGCGCAGGCGGTGA	1020	

Db 961 GCCTGCTGGGCAACACGAGCTGGAGCCGCTTGCTGCAACTCTCTCAGGCCAGGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTTGCCGCTCCTCAGTGGTGAAGAAACACACCTTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGCCGCTCCTCAGTGGTGAAGAAACACACCTTGGCGCCCAATA 1080
Qy 1081 GCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGAGCTGGCAGACAGGTTT 1140
Db 1081 GCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGAGCTGGCAGACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAAGCGAAACCGAATTAATGTAGTACTCACTCAATTAAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAAGCGAAACCGAATTAATGTAGTACTCACTCAATTAAG 1200
Qy 1201 GCACAAATCTCATGTTTGCACAGCTTATCATGACCTGACGCTGACCAATGCTTCTGCG 1260
Db 1201 GCACAAATCTCATGTTTGCACAGCTTATCATGACCTGACGCTGACCAATGCTTCTGCG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGGTATGCTGTGAGGTCGTAATCACTGCATAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGCTGTGAGGTCGTAATCACTGCATAATTCG 1320
Qy 1321 TGTGCTCAAGCGGCACTCCCGTCTGGAATATGTTTTCGCGGAGCATCAACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTCTGGAATATGTTTTCGCGGAGCATCAACGGTT 1380
Qy 1381 CTGGCAAAATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATCGTGGGA 1440
Db 1381 CTGGCAAAATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATCGTGGGA 1440
Qy 1441 ATTGTAGCGGATTAACAAATTTACACAGGAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Db 1441 ATTGTAGCGGATTAACAAATTTACACAGGAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Qy 1501 GCATTTACCAACAGGACCATAGATTAATGAACCTGAGAGGTAATCGTAACTCG 1560
Db 1501 GCATTTACCAACAGGACCATAGATTAATGAACCTGAGAGGTAATCGTAACTCG 1560
Qy 1561 ATTAACCGCGATAAGGCTATAACGGTCTCGCTGAAGTTCGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACCGCGATAAGGCTATAACGGTCTCGCTGAAGTTCGTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTACCGTTGAGCATCCGATTAATCTGGAAGAGAAATTCGCCACAGTT 1680
Db 1621 ACCGGAATTAAGTACCGTTGAGCATCCGATTAATCTGGAAGAGAAATTCGCCACAGTT 1680
Qy 1681 GCGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTGGTGGCTAC 1740
Db 1681 GCGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGTGAATCAACCGGACAAAGCGTTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGTGAATCAACCGGACAAAGCGTTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGTACGTTTCAACGGCAAGCTGTTTACCAGGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGTACGTTTCAACGGCAAGCTGTTTACCAGGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTAACAAGATCTGCTGCGCAACCGCGCAAAACCTCGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTAACAAGATCTGCTGCGCAACCGCGCAAAACCTCGGAA 1920
Qy 1921 GAGATCCGGGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGGGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGCGGGGTATGCTGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGCGGGGTATGCTGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTAAGAGCGTGGGCTGGATTAACGCTGCGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGAGCGTGGGCTGGATTAACGCTGCGCGGAAAGCG 2100

Qy 2101 GGTCTGACCTTCTCTGTTGACCTGATTAATAAAACAAACACATGATGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTAATAAAACAAACACATGATGAGACACCGATTAC 2160
Qy 2161 TCCATCCGAGAAGCTGCTTTTAAATAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCCGAGAAGCTGCTTTTAAATAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCCAAACATCGACACACGCAAAAGTGAATTTATGGTGAACGGTACTTGCACACCTTC 2280
Db 2221 GCATGGTCCAAACATCGACACACGCAAAAGTGAATTTATGGTGAACGGTACTTGCACACCTTC 2280
Qy 2281 AAGGGTCAACCATCCAAACCGTTGCTGGCGTGTGAGCGCAGGTATTAACGGCCCGAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTGCTGGCGTGTGAGCGCAGGTATTAACGGCCCGAGT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTTCTCGAAAACTATCTGCTGATGAGGTTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTTCTCGAAAACTATCTGCTGATGAGGTTCTG 2400
Qy 2401 GAAGCGGTTTAAAGACAAACCCGCTGGTGGTGGCTAGCGTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTTAAAGACAAACCCGCTGGTGGTGGCTAGCGTGAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGGCGAAAGATCCAGTATTTGCGCCACCATCGAAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTTGCGCCACCATCGAAACCGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAAACATCCCGCAGATGTCGGCTTTCTGGTATGCGTGGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAAACATCCCGCAGATGTCGGCTTTCTGGTATGCGTGGTACTGCGGTGATCAACGCC 2580
Qy 2581 GCAGCGCTGCTGAGACTGTGATGAAGCCCTGAAAGCGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCAGCGCTGCTGAGACTGTGATGAAGCCCTGAAAGCGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACAAACAAACAAATTAACAAATTAACAAACCTCGGATCGAGGAGGATTTTCAGAAATTC 2700
Db 2641 AACAAACAAACAAATTAACAAATTAACAAACCTCGGATCGAGGAGGATTTTCAGAAATTC 2700
Qy 2701 GGATCTCT 2709
Db 2701 GGATCTCT 2709

RESULT 9
LOCUS A91965 7475 bp DNA circular PAT 22-JAN-2000
DEFINITION Sequence 1 from Patent WO9822496.
ACCESSION A91965
VERSION A91965.1 GI:6740811
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 7475)
AUTHORS Attwood, M.R. and Hurst, D.N.
TITLE ANTIVIRAL PEPTIDE DERIVATIVES
JOURNAL Patent: WO 9822496-A 1 28-MAY-1998;
HOFFMANN LA ROCHE (CH)
FEATURES
source
1. .7475
Location/Qualifiers
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 81.8%; Score 2700; DB 6; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGCATAAACCTTTCCGGGTATGCGGTATGATAGCGCCCGAAGAGA 60

[illegible]

1081	CGCAAA	CCGCCCTCTCCCGCGCGTGTGCCGATTCA	TTAATGCA	GTGGCAGCA	CGGTTT	1141												
QY	DB	1141	CCC	GACT	TGGAAGCGGCGAGTGA	CGCCAA	CGCAATTA	ATGTGA	TTAGTCA	CTCATTAG	1200							
			1141	CCC	GACT	TGGAAGCGGCGAGTGA	CGCCAA	CGCAATTA	ATGTGA	TTAGTCA	CTCATTAG	1200						
			1141	CCC	GACT	TGGAAGCGGCGAGTGA	CGCCAA	CGCAATTA	ATGTGA	TTAGTCA	CTCATTAG	1200						
QY	DB	1201	GCA	CAATTC	ATCATGTTTGA	CAGCTTAT	TCATCG	ACTGCA	CGGTGC	ACCAATGCTTCTGGCG	1260							
			1201	GCA	CAATTC	ATCATGTTTGA	CAGCTTAT	TCATCG	ACTGCA	CGGTGC	ACCAATGCTTCTGGCG	1260						
			1261	TCAG	CGAGCCAT	CGGAAGCTGTGGTAT	TGGCTG	TCGAGT	CGTAAAT	CACATG	CAATATTCG	1320						
QY	DB	1261	TCAG	CGAGCCAT	CGGAAGCTGTGGTAT	TGGCTG	TCGAGT	CGTAAAT	CACATG	CAATATTCG	1320							
			1321	TGT	CGCTCA	AGCGCACTCCCGTTCT	CGATAAT	TGTTTTT	TTCGCGCCG	ACATCATAA	CGGTT	1380						
			1321	TGT	CGCTCA	AGCGCACTCCCGTTCT	CGATAAT	TGTTTTT	TTCGCGCCG	ACATCATAA	CGGTT	1380						
QY	DB	1381	CTG	CAAAAT	ATTCGAAAT	GAGCTGT	TGCAATTA	TCAT	TCGGCT	CGTATAAT	TGTGGA	1440						
			1381	CTG	CAAAAT	ATTCGAAAT	GAGCTGT	TGCAATTA	TCAT	TCGGCT	CGTATAAT	TGTGGA	1440					
			1441	ATT	GTGAGCGG	ATAACAA	TTTTCAC	ACAGGAAA	CAGCCAGT	CCGTTT	AGGTGTTTT	CACGA	1500					
QY	DB	1441	ATT	GTGAGCGG	ATAACAA	TTTTCAC	ACAGGAAA	CAGCCAGT	CCGTTT	AGGTGTTTT	CACGA	1500						
			1501	GC	ACTTC	CAACAA	GAAGACCATAG	ATTATG	AAATCT	GAAGA	GTAAAT	TGGTAA	CTCG	1560				
			1501	GC	ACTTC	CAACAA	GAAGACCATAG	ATTATG	AAATCT	GAAGA	GTAAAT	TGGTAA	CTCG	1560				
QY	DB	1561	ATT	AACGG	CGATAA	AGGCTATA	ACGGTCT	CGCTG	AAGT	TCGGTAA	GAATTCG	AGAAA	1620					
			1561	ATT	AACGG	CGATAA	AGGCTATA	ACGGTCT	CGCTG	AAGT	TCGGTAA	GAATTCG	AGAAA	1620				
			1621	ACC	GGAA	TTAAAGT	CACGGT	TGACAT	CCCGATA	AACTG	GAAG	AGAAA	TTCCCA	1680				
QY	DB	1621	ACC	GGAA	TTAAAGT	CACGGT	TGACAT	CCCGATA	AACTG	GAAG	AGAAA	TTCCCA	1680					
			1681	CG	CAACT	GTGGCGT	GCCCTG	CAATTA	TCTTCT	TGGG	CAACA	CGACCG	CTTTGGTGG	CTAC	1740			
			1681	CG	CAACT	GTGGCGT	GCCCTG	CAATTA	TCTTCT	TGGG	CAACA	CGACCG	CTTTGGTGG	CTAC	1740			
QY	DB	1741	GCT	CAATCT	GTGCC	CTGT	TGGTGA	AAATCA	CCCCGG	ACAAAG	CGTTC	CCAGG	CAACAG	CTGTAT	1800			
			1741	GCT	CAATCT	GTGCC	CTGT	TGGTGA	AAATCA	CCCCGG	ACAAAG	CGTTC	CCAGG	CAACAG	CTGTAT	1800		
			1801	CCG	TTTAC	TCGGAT	TGCCGT	ACGTTA	CAAC	CGGC	ACAGTGA	TTGCTTA	ACCGAT	CGCTG	TTT	1860		
QY	DB	1801	CCG	TTTAC	TCGGAT	TGCCGT	ACGTTA	CAAC	CGGC	ACAGTGA	TTGCTTA	ACCGAT	CGCTG	TTT	1860			
			1861	GA	CGGTTAT	CGCTG	ATTATA	CA	AAAGAT	TCGCTG	CCGAC	CCCGC	CAAAA	AACTGG	GAA	1920		
			1861	GA	CGGTTAT	CGCTG	ATTATA	CA	AAAGAT	TCGCTG	CCGAC	CCCGC	CAAAA	AACTGG	GAA	1920		
QY	DB	1921	GAG	AT	CCCGCG	CTGGATA	AAAGAA	CTGAA	AGCG	CAAAAG	GTAAAG	CGCGCT	GTAT	TTCA	1980			
			1921	GAG	AT	CCCGCG	CTGGATA	AAAGAA	CTGAA	AGCG	CAAAAG	GTAAAG	CGCGCT	GTAT	TTCA	1980		
			1981	CTG	CAAGAA	CCGTAT	CTCA	CTGG	CCGCTGA	TTGCTG	CTG	ACGGGG	GTTTAT	TCGTT	CAAG	2040		
QY	DB	1981	CTG	CAAGAA	CCGTAT	CTCA	CTGG	CCGCTGA	TTGCTG	CTG	ACGGGG	GTTTAT	TCGTT	CAAG	2040			
			2041	TAT	GA	AAACG	CAAGTAC	GCATTA	AAAGAC	GTGG	CGTGA	TACG	CTG	CGCG	CAAG	2100		
			2041	TAT	GA	AAACG	CAAGTAC	GCATTA	AAAGAC	GTGG	CGTGA	TACG	CTG	CGCG	CAAG	2100		
QY	DB	2101	GGT	CTG	ACCTT	CTCTGG	TTGAC	CTTAT	TA	AAAAA	CAAA	CA	CAATGA	TG	CAGACA	CCGATTAC	2160	
			2101	GGT	CTG	ACCTT	CTCTGG	TTGAC	CTTAT	TA	AAAAA	CAAA	CA	CAATGA	TG	CAGACA	CCGATTAC	2160
			2161	TCC	AT	TCG	CAGAA	GTG	CTTTT	ATAA	AGG	CGAAA	CAG	CGAT	TG	ACCAT	TC	AA

```
QY 2221 GCATGGTCCAAATCGACACACGAAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACACGAAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGTCACACCATCAAAACCGTTCTTGGCGTGTGAGCGAGGTATTAACCGCGCCAGT 2340
Db 2281 AAGGTCACACCATCAAAACCGTTCTTGGCGTGTGAGCGAGGTATTAACCGCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTAAATAAGACAAACCGTGGTGCCTGAGCGTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTAAATAAGACAAACCGTGGTGCCTGAGCGTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCACGATTTGCGCCACCATGGAAGCCCGAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTTGCGCCACCATGGAAGCCCGAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCGCGAGATGTCGCTTTCTGCTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCGCGAGATGTCGCTTTCTGCTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTGTGAGTGTGATGAAGCCCTGAAAGCGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTGTGAGTGTGATGAAGCCCTGAAAGCGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAACAATAACAACAACCTCGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACAACCTCGGATCGAGGGAAGGATTCAGAAATTC 2700

RESULT 10
AR031992 LOCUS AR031992 7475 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5866684.
ACCESSION AR031992
VERSION AR031992.1 GI:5946281
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 7475)
Attwood,M.Richard., Hurst,D.Nigel., Jones,P.Stephen.,
Kay,P.Brittain., Raynham,T.Michael. and Wilson,F.Xavier.
TITLE Peptidyl inhibitors of viral proteases
JOURNAL Patent: US 5866684-A 1 02-FEB-1999;
FEATURES Location/Qualifiers
source 1..7475
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 81.8%; Score 2700; DB 6; Length 7475;
Best Local Similarity 100.0%; Pred.No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGAACAACTTTCCGCGTATGGATGATAGCGCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGAACAACTTTCCGCGTATGGATGATAGCGCCGGAAGAGA 60
QY 61 GTCAATTCAGGCGGTGAATGTGAACAGTAACTGATAGTATCGATGTCGAGAGTATCCG 120
Db 61 GTCAATTCAGGCGGTGAATGTGAACAGTAACTGATAGTATCGATGTCGAGAGTATCCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCCGCGTGTGGAACAGGCGGAGCCAGCGTTTCTCGAAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCCGCGTGTGGAACAGGCGGAGCCAGCGTTTCTCGAAAA 180
QY 181 CGCGGAAAAAGTGAAGCGCGATGCGGAGTGAATTAATTCCTCCAAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGCGATGCGGAGTGAATTAATTCCTCCAAACCGCGTGGCAC 240
```

```
Db 181 CGCGGAAAAAGTGAAGCGCGATGCGGAGTGAATTAATTCCTCCAAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGCAAAACAGTCTGTTGCTGATTGGCGTTGGCACTCAGTCTGCGCCCTGC 300
Db 241 AACAACTGGCGGCAAAACAGTCTGTTGCTGATTGGCGTTGGCACTCAGTCTGCGCCCTGC 300
QY 301 ACGCGCGCTCGCAAAATTTGTCGCGCGGATTAATCTCGCGCGGATCAACTGGTGGCCAGCG 360
Db 301 ACGCGCGCTCGCAAAATTTGTCGCGCGGATTAATCTCGCGCGGATCAACTGGTGGCCAGCG 360
QY 361 TGTGTGTGTCGATGTAGAAACGAAGCGGCTGAAAGCTCTGAAAGCGCGGTGCACAATC 420
Db 361 TGTGTGTGTCGATGTAGAAACGAAGCGGCTGAAAGCTCTGAAAGCGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACCAAGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACCAAGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAAATGTTCCGGCGTTATTTCTTTGATGTCTCTGACCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAAATGTTCCGGCGTTATTTCTTTGATGTCTCTGACCAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGCTACGCGCTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGCTACGCGCTGGGCGTGGAGCATC 600
QY 601 TGGTCGCAATTTGGGTCCACGACAAATCGCGTGTAGCGGCGCCATTAAGTTCTGTCGCG 660
Db 601 TGGTCGCAATTTGGGTCCACGACAAATCGCGTGTAGCGGCGCCATTAAGTTCTGTCGCG 660
QY 661 CGCGTCTGCTGTGGCTGGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCTGTGGCTGGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAAACGGGAAGCGGCTGAGTGCCATGTCGCGTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAAACGGGAAGCGGCTGAGTGCCATGTCGCGTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGCGCATGCTTCCACATCGCATGCTGTTGCCAACGATCAGATGGCGCTCGGCGCAA 840
Db 781 ATGAGGCGCATGCTTCCACATCGCATGCTGTTGCCAACGATCAGATGGCGCTCGGCGCAA 840
QY 841 TGC CGCGCCATTAACGAGTCCGGCTCGCGTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TGC CGCGCCATTAACGAGTCCGGCTCGCGTGGTGGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGGAAGACGCTCATGTTATATCCCGCGTTAACCCACCATCAACAGGATTTTC 960
Db 901 ACGATACCGGAAGACGCTCATGTTATATCCCGCGTTAACCCACCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGGCAAAACGAGCTGGAGCCGTTGCTGCAACTCTCTCAGGGCGAGGCGGTGA 1020
Db 961 GCCTGCTGGGGCAAAACGAGCTGGAGCCGTTGCTGCAACTCTCTCAGGGCGAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCGCTCTCATCTGGTGAAGAAAGAAACACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCGCTCTCATCTGGTGAAGAAAGAAACACCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGCGTTGCGCGATTTATTATGACGCTGGCAACAGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGCGTTGCGCGATTTATTATGACGCTGGCAACAGAGTTT 1140
QY 1141 CCGCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCTCACTATTAG 1200
Db 1141 CCGCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCTCACTATTAG 1200
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGCGAGCCATCGGAAGCTGTGATGCTGTGAGGTCGTAATAATCACTGCAATAATTCG 1320
Db 1261 TCAGGCGAGCCATCGGAAGCTGTGATGCTGTGAGGTCGTAATAATCACTGCAATAATTCG 1320
```

```
QY 1321 TGTGCTCAAGCGCAGCTCCCGTTCCTGGATAATGTTTTTTTGGCCGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGCAGCTCCCGTTCCTGGATAATGTTTTTTTGGCCGACATCATACGGTT 1380
QY 1381 CTGGCAATATCTGAATAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAATATCTGAATAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
QY 1441 ATTGTGAGCGGATAACAATTTTACACAGAAACAGCCAGTCCGTTTAGTGTTTTTCA 1500
Db 1441 ATTGTGAGCGGATAACAATTTTACACAGAAACAGCCAGTCCGTTTAGTGTTTTTCA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATATGAAAACTGAAGAGGTAACCTGTTAACTCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATATGAAAACTGAAGAGGTAACCTGTTAACTCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATTAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATTAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAACTGGAGAGAAAATTCACACAGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAACTGGAGAGAAAATTCACACAGTT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTTGGCTGAATATCAACCGGACAAAGCGTTCCAGACAACTGTAT 1800
Db 1741 GCTCAATCTGGCTGTTGGCTGAATATCAACCGGACAAAGCGTTCCAGACAACTGTAT 1800
QY 1801 CGTTTTACTCGGGATGCGGTAGCTTACAAAGCAAGCTGATGCTTACCCGATCGCTGT 1860
Db 1801 CGTTTTACTCGGGATGCGGTAGCTTACAAAGCAAGCTGATGCTTACCCGATCGCTGT 1860
QY 1861 GAAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGAA 1920
Db 1861 GAAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGAA 1920
QY 1921 GAGATCCGGCGCTGGATAAGAACTGAAGCGAAAGGTGAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGGCGCTGGATAAGAACTGAAGCGAAAGGTGAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCAACCTGGCGCTGATGCTGCTGACGGGGTTATCGCTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCAACCTGGCGCTGATGCTGCTGACGGGGTTATCGCTTCAAG 2040
QY 2041 TATGAAAAACGGCAAGTACGACATTAAGACGTTGGCGTGGATAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAAACGGCAAGTACGACATTAAGACGTTGGCGTGGATAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTGGTGAACCTGATTAATAAACAACACATGAATGCGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTGGTGAACCTGATTAATAAACAACACATGAATGCGAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCCCTTTAATAAGCGGAAACAGCGATCACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCCCTTTAATAAGCGGAAACAGCGATCACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCAAATCGACACCAAGCAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTTC 2280
Db 2221 GCATGGTCCAAATCGACACCAAGCAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTTGGTGGCTGCTGAGCGCAGGTAATTAACCGCGCCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTTGGTGGCTGCTGAGCGCAGGTAATTAACCGCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAACTATCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAACTATCTGACTGATGAAGGTCTG 2400
```

```
QY 2401 GAAAGCGTTAATAAAGACAAACCCGCTGGTCCGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
Db 2401 GAAAGCGTTAATAAAGACAAACCCGCTGGTCCGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
QY 2461 TTGGGGAAGATCCACGTTATTTGCGGCCACCATGGAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGGGAAGATCCACGTTATTTGCGGCCACCATGGAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCGCTTTTCTGGTATGCGCTGCTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCGCTTTTCTGGTATGCGCTGCTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCCGACACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCCGACACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAACAATAACAACAACTCGGATCGAGGGAAGATTTCAGAAATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACAACTCGGATCGAGGGAAGATTTCAGAAATTC 2700

RESULT 11
AR207294
LOCUS AR207294 7475 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 1 from patent US 6372883.
ACCESSION AR207294
VERSION AR207294.1 GI:21506162
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7475)
AUTHORS Attwood,M.Richard., Hurst,D.Nigel., Jones,P.Stephen.,
Kay,P.Brittain., Raynham,T.Michael. and Wilson,P.Xavier.
TITLE Antiviral medicaments
JOURNAL Patent: US 6372883-A 1 16-APR-2002;
FEATURES
source Location/Qualifiers
1..7475
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 81.8%; Score 2700; DB 6; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGTATGCGATGATAGCGCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGTATGCGATGATAGCGCCGGAAGAGA 60
QY 61 GTCATTCAGGGTGGTGAATGTGAAACCAAGTAAACGTTATACGATGTCGACAGTATGCCG 120
Db 61 GTCATTCAGGGTGGTGAATGTGAAACCAAGTAAACGTTATACGATGTCGACAGTATGCCG 120
QY 121 GTGTCCTTATCAGACCGCTTTCCCGCGTGGTGAACACAGCCACCAACGTTTCTGCGAAAA 180
Db 121 GTGTCCTTATCAGACCGCTTTCCCGCGTGGTGAACACAGCCACCAACGTTTCTGCGAAAA 180
QY 181 CGCGGAAAAAGTGAAGCGGCGATGCGGAGCTGAATTAATTCCTCCACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGGCGATGCGGAGCTGAATTAATTCCTCCACCGCGTGGCAC 240
QY 241 AACAACTCGCGGGCAACACAGTCTGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTTCG 300
Db 241 AACAACTCGCGGGCAACACAGTCTGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTTCG 300
QY 301 ACGCCCGCTCGCAAAATTTCTCGCGCGATTAATAATCTCGCCCGATCAATGGGTGCGACG 360
Db 301 ACGCCCGCTCGCAAAATTTCTCGCGCGATTAATAATCTCGCCCGATCAATGGGTGCGACG 360
QY 361 TGGTGGTTCGATGTTAGAAAGAACCGCGCTCGAAGCCCTGTAAAGCGGCGGTGCAATTC 420
Db 361 TGGTGGTTCGATGTTAGAAAGAACCGCGCTCGAAGCCCTGTAAAGCGGCGGTGCAATTC 420
```


Db 2581 GCGCGGTGTCAGACTGTGATGAAGCCCTGGAAGACGCCGACATAATTCGAGCTCG 2640

Qy 2641 AACAAACAACAATAACAATAACAACACCTCGGGATCGAGGGAAGGATTTCAGAAATTC 2700

Db 2641 AACAAACAACAATAACAATAACAACACCTCGGGATCGAGGGAAGGATTTCAGAAATTC 2700

RESULT 12

AX377532

LOCUS AX377532 9191 bp DNA linear PAT 18-MAR-2002

DEFINITION Sequence 9 from Patent WO0212553.

ACCESSION AX377532

VERSION AX377532.1 GI:19573718

KEYWORDS .

SOURCE synthetic construct

ORGANISM synthetic construct

other sequences; artificial sequences.

REFERENCE 1

AUTHORS Kappel,A., Polakowski,T., Pignot,M., Windhab,N., Behrensdoerf,H. and Muth,J.

TITLE Method for detecting mutations in nucleotide sequences

JOURNAL Patent: WO 0212553-A 9 14-FEB-2002;

Nanogen Recognomics GmbH (DE)

FEATURES

source

1. .9191

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

/note="pMALc2x-mutS Plasmid"

ORIGIN

Query Match 81.8%; Score 2698; DB 6; Length 9191;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 2701; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGCAGAAACCTTTCGGCGGTATGGCATGATAGGCCCGGAAGAGA 60

Db 1 CCGACACCATCGAATGGTGCAGAAACCTTTCGGCGGTATGGCATGATAGGCCCGGAAGAGA 60

Qy 61 GTCATTCAGGGTGTGAATGTGAACACGATGAACAGTATACGATCTCGCAGAGTATGCCG 120

Db 61 GTCATTCAGGGTGTGAATGTGAACACGATGAACAGTATACGATCTCGCAGAGTATGCCG 120

Qy 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACAGGCCACGACGTTTCTGCGAAAA 180

Db 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACAGGCCACGACGTTTCTGCGAAAA 180

Qy 181 CCGGGAAAAAGTGAAGCGCGATGGCGAGCTGAAATTACATTCCCAACCGCGTGGCAC 240

Db 181 CCGGGAAAAAGTGAAGCGCGATGGCGAGCTGAAATTACATTCCCAACCGCGTGGCAC 240

Qy 241 AACAACTGGCGGCAACAGTCTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300

Db 241 AACAACTGGCGGCAACAGTCTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300

Qy 301 AGCGCCGTGCGCAAAATTTGCGCGCGATTAATAATCTCGCGCGATCAACTGGGTGCGAGCG 360

Db 301 AGCGCCGTGCGCAAAATTTGCGCGCGATTAATAATCTCGCGCGATCAACTGGGTGCGAGCG 360

Qy 361 TGGTGGTGTGATGTAGAACGAAAGCGCGTGCAGACCTGTAAAGCGGCGGTGCACAATC 420

Db 361 TGGTGGTGTGATGTAGAACGAAAGCGCGTGCAGACCTGTAAAGCGGCGGTGCACAATC 420

Qy 421 TTCTCGCGCAACGCTCAGTGGGTGATCAATTAATCTATCCGCTGGATGACAGGATGCCA 480

Db 421 TTCTCGCGCAACGCTCAGTGGGTGATCAATTAATCTATCCGCTGGATGACAGGATGCCA 480

Qy 481 TTGCTGTGGAAGCTGCTGCAATAATGTTCCGGCGTTATTTCTGATGTCTCTGACCCAGA 540

Db 481 TTGCTGTGGAAGCTGCTGCAATAATGTTCCGGCGTTATTTCTGATGTCTCTGACCCAGA 540

Qy 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600

Db 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600

Qy 601 TGGTGCATTTGGGTACACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTTCTGTCTCGG 660

Db 601 TGGTGCATTTGGGTACACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTTCTGTCTCGG 660

Qy 661 CGGCTGCGCTGCGTGGCGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720

Db 661 CGGCTGCGCTGCGTGGCGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720

Qy 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAACCAATGCAAAATGCTGA 780

Db 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAACCAATGCAAAATGCTGA 780

Qy 781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGGTGGCAACGATCAGATGGCGCGCAA 840

Db 781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGGTGGCAACGATCAGATGGCGCGCAA 840

Qy 841 TGGCGCCATTTACCGAGTCCGGGCTGCGGCTGCGGATATCTCGGTAGTGGGATACG 900

Db 841 TGGCGCCATTTACCGAGTCCGGGCTGCGGCTGCGGATATCTCGGTAGTGGGATACG 900

Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAACAGGATTTTC 960

Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAACAGGATTTTC 960

Qy 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGTCGAACCTCTCTCAGGCGCCAGGCGTGA 1020

Db 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGTCGAACCTCTCTCAGGCGCCAGGCGTGA 1020

Qy 1021 AGGGCAATCAGCTGTGTCCTCTCACTGTTGAAAGAAAAACCACTGGCGCCCAATA 1080

Db 1021 AGGGCAATCAGCTGTGTCCTCTCACTGTTGAAAGAAAAACCACTGGCGCCCAATA 1080

Qy 1081 CGCAAAACCGCTCTCCCGCGCTGGCGGATCAATTAATCGAGTGGCGACGACAGGTTT 1140

Db 1081 CGCAAAACCGCTCTCCCGCGCTGGCGGATCAATTAATCGAGTGGCGACGACAGGTTT 1140

Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200

Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200

Qy 1201 GCACAATTTCTCATGTTGACAGCTTATCATCGACTGACGGTGCACCAATGCTTCTGGCG 1260

Db 1201 GCACAATTTCTCATGTTGACAGCTTATCATCGACTGACGGTGCACCAATGCTTCTGGCG 1260

Qy 1261 TCAGGCAGCCATCGGAAGCTGTGTTAGCTGTGCGAGTCTGTAATCACTGCATTAATTCG 1320

Db 1261 TCAGGCAGCCATCGGAAGCTGTGTTAGCTGTGCGAGTCTGTAATCACTGCATTAATTCG 1320

Qy 1321 TGTGCTCAAGCGGCACCTCCCGTTCTGGATTAATGTTTTTGGCGCGACATCATAAACGGTT 1380

Db 1321 TGTGCTCAAGCGGCACCTCCCGTTCTGGATTAATGTTTTTGGCGCGACATCATAAACGGTT 1380

Qy 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGTGGA 1440

Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGTGGA 1440

Qy 1441 ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500

Db 1441 ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500

Qy 1501 GCACCTTACCAACAGGACCATAGATTATGAAACTGAAAGAGGTAAACTGGTAATCTCGG 1560

Db 1501 GCACCTTACCAACAGGACCATAGATTATGAAACTGAAAGAGGTAAACTGGTAATCTCGG 1560

Qy 1561 ATTTAAACGGCGATAAAGGCTATAACGGCTCTCGCTGAAGTTCGTTAAAGAAATTCGAGAAAGAT 1620

Db 1561 ATTTAAACGGCGATAAAGGCTATAACGGCTCTCGCTGAAGTTCGTTAAAGAAATTCGAGAAAGAT 1620

Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCACAGGTTT 1680

D _b	1621	ACGGAA	TTAAAG	TCA	CGTTGAG	CAT	CCGGATA	AACTCG	GAAGAGAA	AT	CCCCAG	TT	1690		
Q _y	1681	GCGCA	ATGGCG	ATGGCC	TGACAT	TA	CTTCTGG	GCACAC	GACGAC	CGCTTT	GGTG	GGCTAC	1740		
D _b	1681	GCGCA	ATGGCG	ATGGCC	TGACAT	TA	CTTCTGG	GCACAC	GACGAC	CGCTTT	GGTG	GGCTAC	1740		
Q _y	1741	GCTCA	ATCTGG	CTGTTGG	CTGA	AAATCA	CCCCG	GA	AAAGCG	TTCC	AGCA	AACTGTAT	1800		
D _b	1741	GCTCA	ATCTGG	CTGTTGG	CTGA	AAATCA	CCCCG	GA	AAAGCG	TTCC	AGCA	AACTGTAT	1800		
Q _y	1801	CCGTT	TAC	CTGG	ATG	CGT	TGTTA	CA	ACGCA	AGCTGA	TGTTCT	ACCGAT	CGCTGTT	1860	
D _b	1801	CCGTT	TAC	CTGG	ATG	CGT	TGTTA	CA	ACGCA	AGCTGA	TGTTCT	ACCGAT	CGCTGTT	1860	
Q _y	1861	GAAG	CGTTAT	CGCTGA	TTTAA	CA	AAAGAT	CTG	TGCGGA	ACCCG	CA	AAACCTGG	GA	1920	
D _b	1861	GAAG	CGTTAT	CGCTGA	TTTAA	CA	AAAGAT	CTG	TGCGGA	ACCCG	CA	AAACCTGG	GA	1920	
Q _y	1921	GAGAT	CCCG	CGCTGG	ATA	AGAACTGA	AAAGCG	AAAGTA	AGCGG	CGCTGA	TGTTCA	AC	1980		
D _b	1921	GAGAT	CCCG	CGCTGG	ATA	AGAACTGA	AAAGCG	AAAGTA	AGCGG	CGCTGA	TGTTCA	AC	1980		
Q _y	1981	CTGCA	AGAAC	CGTACT	CTCAC	CTG	CGCTGA	TGCTG	CTGAC	GCGGGT	TATG	CGTTCA	AG	2040	
D _b	1981	CTGCA	AGAAC	CGTACT	CTCAC	CTG	CGCTGA	TGCTG	CTGAC	GCGGGT	TATG	CGTTCA	AG	2040	
Q _y	2041	TATGA	AAAC	CGCA	GTG	CAAT	TAA	AGAC	GTG	CGCTGG	TGAT	AC	GTGCGG	GA	2100
D _b	2041	TATGA	AAAC	CGCA	GTG	CAAT	TAA	AGAC	GTG	CGCTGG	TGAT	AC	GTGCGG	GA	2100
Q _y	2101	GGTCT	GAC	TTTCT	CGTGGT	TGAC	CTGAT	TAA	AAACAA	CACAT	GAAT	CAGAC	ACCGAT	TAC	2160
D _b	2101	GGTCT	GAC	TTTCT	CGTGGT	TGAC	CTGAT	TAA	AAACAA	CACAT	GAAT	CAGAC	ACCGAT	TAC	2160
Q _y	2161	TCCAT	CGCA	AGCTG	CTTTAA	TAA	AGCG	AAACAG	CGAT	TGAC	CAAT	CAACG	CCGTG	G	2220
D _b	2161	TCCAT	CGCA	AGCTG	CTTTAA	TAA	AGCG	AAACAG	CGAT	TGAC	CAAT	CAACG	CCGTG	G	2220
Q _y	2221	GCAT	GTG	TCCAA	CTCG	ACAC	CAG	CAAA	AGTGA	ATTAT	TGTTG	TGAC	GTGCG	CACTTC	2280
D _b	2221	GCAT	GTG	TCCAA	CTCG	ACAC	CAG	CAAA	AGTGA	ATTAT	TGTTG	TGAC	GTGCG	CACTTC	2280
Q _y	2281	AAGGT	CA	CCAT	CCAA	CCGTTG	CTG	TGGG	TGCTG	AGCG	CAG	GTATTA	ACGCG	CCAGT	2340
D _b	2281	AAGGT	CA	CCAT	CCAA	CCGTTG	CTG	TGGG	TGCTG	AGCG	CAG	GTATTA	ACGCG	CCAGT	2340
Q _y	2341	CCGAA	CA	AAAG	CTGG	CA	AAAG	CTTC	CGA	AACTAT	CTG	CTG	AGTGA	AGTCTG	2400
D _b	2341	CCGAA	CA	AAAG	CTGG	CA	AAAG	CTTC	CGA	AACTAT	CTG	CTG	AGTGA	AGTCTG	2400
Q _y	2401	GAAG	CGGTTT	ATA	AAAG	CAAA	CCGCTGG	TG	CGCTG	AGCTGA	AGTCTT	AC	GAGGA	AG	2460
D _b	2401	GAAG	CGGTTT	ATA	AAAG	CAAA	CCGCTGG	TG	CGCTG	AGCTGA	AGTCTT	AC	GAGGA	AG	2460
Q _y	2461	TTGCG	GA	AAAG	ATCCA	CGT	TATG	CGCC	ACCAT	GG	AAACG	CCAG	AAAGTGA	ATCATG	2520
D _b	2461	TTGCG	GA	AAAG	ATCCA	CGT	TATG	CGCC	ACCAT	GG	AAACG	CCAG	AAAGTGA	ATCATG	2520
Q _y	2521	CCGA	ACAT	CCCG	AGATG	TG	CTGG	TATG	CGCGT	TGCTG	AGTACT	CGG	TGATCA	ACGCG	2580
D _b	2521	CCGA	ACAT	CCCG	AGATG	TG	CTGG	TATG	CGCGT	TGCTG	AGTACT	CGG	TGATCA	ACGCG	2580
Q _y	2581	GCCAG	CGT	CTG	AGTCTG	ATGA	AGCCCTG	AA	AGAC	CGCG	CAG	ATTA	TTG	AGCTCG	2640
D _b	2581	GCCAG	CGT	CTG	AGTCTG	ATGA	AGCCCTG	AA	AGAC	CGCG	CAG	ATTA	TTG	AGCTCG	2640
Q _y	2641	AACA	CAAC	AAC	AAAT	TAA	CAAT	TAA	CAAC	AACTCG	GGAT	CGAGG	GAAGT	TTT	2700
D _b	2641	AACA	CAAC	AAC	AAAT	TAA	CAAT	TAA	CAAC	AACTCG	GGAT	CGAGG	GAAGT	TTT	2700
Q _y	2701	GGAT	CC												2706
D _b	2701	GGAT	CC												2706

```

RESULT 13
AF097412      6724 bp      DNA      circular SYN 28-JAN-2002
LOCUS
DEFINITION   Expression vector pMBP-parallel1, complete sequence.
ACCESSION   AF097412
VERSION      AF097412.1 GI:3983120
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 6724)
Sheffield,P., Garrard,S. and Derewenda,Z.
Overcoming expression and purification problems of RhoGDI using a
family of 'parallel' expression vectors
Protein Expr. Purif. 15 (1), 34-39 (1999)
99150479
PUBMED
10024467
REFERENCE
2 (bases 1 to 6724)
Sheffield,P.J., Garrard,S.M. and Derewenda,Z.S.
Direct Submission
Submitted (05-OCT-1998) Molecular Physiology and Biological
Physics, University of Virginia, 4215 Jordan Hall, 1300 Jefferson
Park Avenue, Charlottesville, Virginia 22908, USA
JOURNAL
FEATURES
source
1..6724
/organism="Expression vector pMBP-parallel1"
/mol_type="other DNA"
/db_xref="taxon:84479"
204..1163
/gene="lacIq"
204..1163
/gene="lacIq"
/codon_start=1
/transl_table=1
/product="lac repressor"
/protein_id="AAC83812.1"
/db_xref="GI:3983121"
/translation="WAELENYIPNRVAQQLAGKQSLIGVATSSSLAHAPSQIVAIAKIS
RADQAGSVVMSVRSVEACKAAVHNLLAQVRSGLIINYPDDDAIAVEAACTNY
PALFDVDSQDPINSLIFSHEDGTGLGVEHLVAGHQOIALLAGLSSVSRLLRAGW
KHYLRNIOQPIAREGWSAMSQFQOTQMLNEGIVPTAMLVANDQMALGAMRAITE
SGLRGYDTSVGYDDTDSSCYIPPLTTIKQDFRLIGOTSVDRLQLSQGQAVKNGO
LLPVSLVRKVTTLAPNTQTASPRALADSLMQLAQVRSLRSGQ"
1453..1481
/note="tac; inducible promoter"
1528..2994
/gene="male"
1528..2994
/gene="male"
/codon_start=1
/transl_table=1
/product="maltose binding protein"
/protein_id="AAC83813.1"
/db_xref="GI:3983122"
/translation="WKIEEGKLVITWINGDKYNGLAEVGKKFEKDTGIKVTVEHPDKL
BEKFPQVATGDPDIIFWAHDRTFGVAGSGLLAEITPDKAFQDKLYPTWDVAVRYNG
KLIAVPAVEALSLIYNKDLINPPKTKWEELPALDKELKAKGKSALMENLQEPYFTWP
LIADGGVAFYKNGKYDKIDQVDNAGAKGLFLVDLIKNKHNNADTDYSIAEAPF
NKGETAMTINGPAWSNIDTISKVNGVTVLPTFKGQSPKDPFVGLSAGINAASPNKEL
AKFEYENLYLTDGELAVNKDKPLGVALKSYEBELAKDPFIAATMENAAQKEMPNFI
QMGAFWYAVRTAVINAAAGRTQVDEALKDAQTSSNNNNNNNNNNLGIOTTENLYF
QGAMDPEFGKGLRRRAQLVRPLSNLEPAVSRHAVPSLALAVLQRRDWNENPGVTQLNRL
AAHPFPASRNSEEARTRDPSPQQLRSLNGEWQLGCFGG"
2633..2698
/gene="male"
/note="protein spacer region"
2699..2709
/gene="male"
/note="Tobacco etch virus protease cleavage site"
2711..2810
/gene="male"
misc_feature
misc_feature
misc_feature

```

		/note="multiple cloning site"			
gene		3426..4356		Db	721
promoter	/gene="bla"			Qy	781
	/gene="bla"	3426..3454		Db	781
	/gene="bla"	3496..4356		Qy	841
CDS	/gene="bla"			Db	841
	/codon_start=1			Qy	901
	/transl_table=11			Db	901
	/product="beta-lactamase"			Qy	901
	/protein_id="AAC83814.1"			Db	901
	/db_xref="GI:3983123"			Qy	961
	/translation="MSIQHFRVALLIPFAAFLPVFAHPETLVKVKDAEDQLGARVGY			Db	961
	IELDNSKILDSFPEERKPMSTFKVLLDGLAVLSRVADGQEQRIHYSDNDLVE			Qy	1021
	YSPVTEKHLTDGMTVRELCSAAITWSDNTAANLLLTITGGPKELTAFIHNMGDHLRL			Db	1021
	DRWPELNEAIFPNDRDTPVPVAMATTLRLITGELLTLASRQQLIDWMEADKVGAPL			Qy	1081
	LRSLAPGWFTADKSGAGERSGRIIAALPGDKPSRIWVIYTTGSGQATMDERNQIA			Db	1081
	EIGASLIKHW"			Qy	1141
ORIGIN					
Query Match 81.1%; Score 2676.2; DB 12; Length 6724;					
Best Local Similarity 99.9%; Pred. No. 0;					
Matches 2678; Conservative 0; Mismatches 3; Indels 0; Gaps 0;					
Qy	1	CCGACACCATCAATGGTGC	AAACCTTTTCGGGTATGG	CAATAGGCCCGGAAGAGA	60
Db	1	CCGACACCATCAATGGTGC	AAACCTTTTCGGGTATGG	CAATAGGCCCGGAAGAGA	60
Qy	61	GTCAATTACAGGTGTGAAT	GTGAACACAGTAACGTTAT	ACCATGTCGCAGAGTATCGG	120
Db	61	GTCAATTACAGGTGTGAAT	GTGAACACAGTAACGTTAT	ACCATGTCGCAGAGTATCGG	120
Qy	121	GTGTCTCTTATCAGACCG	GTTCCTCCGGTGTGAAC	CCAGCCACGCTTTCTCGG	180
Db	121	GTGTCTCTTATCAGACCG	GTTCCTCCGGTGTGAAC	CCAGCCACGCTTTCTCGG	180
Qy	181	CGCGGAAAGTGAAGCGG	CGATGGCGAGCTGAAAT	TACATTCCTCCACACCG	240
Db	181	CGCGGAAAGTGAAGCGG	CGATGGCGAGCTGAAAT	TACATTCCTCCACACCG	240
Qy	241	AACAACTGCGGGCAAC	AGTCTGTGATTTGGGCT	TGGCCACCTCCAGTCTGC	300
Db	241	AACAACTGCGGGCAAC	AGTCTGTGATTTGGGCT	TGGCCACCTCCAGTCTGC	300
Qy	301	ACGCGCGTTCGCAAAAT	TGTTCGGCGGATTAATCT	CGCGCGATCAACTGGGTGC	360
Db	301	ACGCGCGTTCGCAAAAT	TGTTCGGCGGATTAATCT	CGCGCGATCAACTGGGTGC	360
Qy	361	TGGTGTGTCGATGTAG	AAACGAGCGGCTGAAAG	CGCGGTGCAATC	420
Db	361	TGGTGTGTCGATGTAG	AAACGAGCGGCTGAAAG	CGCGGTGCAATC	420
Qy	421	TTCTCGCGCAACGCTCA	GTGGGCTGATCAATTAAT	CTCCGCTGGATGACCAAG	480
Db	421	TTCTCGCGCAACGCTCA	GTGGGCTGATCAATTAAT	CTCCGCTGGATGACCAAG	480
Qy	481	TTGCTGTGGAAGCTGCT	CGCATTAATGTTCCGGG	CTTATTTCTGATGTTCTG	540
Db	481	TTGCTGTGGAAGCTGCT	CGCATTAATGTTCCGGG	CTTATTTCTGATGTTCTG	540
Qy	541	CACCCATCAACAGTAT	TTTCTCCATGAAGACG	GTAACGCGTGGAGCATC	600
Db	541	CACCCATCAACAGTAT	TTTCTCCATGAAGACG	GTAACGCGTGGAGCATC	600
Qy	601	TGGTGCATTTGGTTCAC	CGAAATTCGGCTGTAG	CGGGCCCATTAAGTCTG	660
Db	601	TGGTGCATTTGGTTCAC	CGAAATTCGGCTGTAG	CGGGCCCATTAAGTCTG	660
Qy	661	CGCGTCTGGCTGGCTG	CGCAATAATCTCACTC	CGCAATCAAAATTCAGC	720
Db	661	CGCGTCTGGCTGGCTG	CGCAATAATCTCACTC	CGCAATCAAAATTCAGC	720
Qy	721	CGGAACGGGAAGCGCA	CTGGAGTGCCATGTCC	GGTTTTCACAAACCATG	780


```
Db      841  TCCGCCCATTACCAGTCCGGGCTGGCGTTGGTGGCGATATCTCGGTAGTGGGATAGC 900
Qy      901  ACGATACCGAAGACAGCTCATGTTATATCCGCCGTTAAACCAACCAACAGGATTTTC 960
Db      901  ACGATACCGAAGACAGCTCATGTTATATCCGCCGTTAAACCAACCAACAGGATTTTC 960
Qy      961  GCCTGCTGGGGCAACACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db      961  GCCTGCTGGGGCAACACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Qy     1021  AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAACCAACCTGGCGGCCAATA 1080
Db     1021  AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAACCAACCTGGCGGCCAATA 1080
Qy     1081  CGCAAAACCGCTCTCCCGCGGTTGGCGATTCAATTAATGCACTGCGACGACAGGTTT 1140
Db     1081  CGCAAAACCGCTCTCCCGCGGTTGGCGATTCAATTAATGCACTGCGACGACAGGTTT 1140
Qy     1141  CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTATTAG 1200
Db     1141  CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTATTAG 1200
Qy     1201  GCACAATCTCATGTTTGACAGCTTATCATCGACTGACGGTGCAACCAATGCTTCTGGCG 1260
Db     1201  GCACAATCTCATGTTTGACAGCTTATCATCGACTGACGGTGCAACCAATGCTTCTGGCG 1260
Qy     1261  TCAGGACGCAATCGGAAGCTGTGGTATGCTGTGAGGTCGTAATCACTGCAATATTCG 1320
Db     1261  TCAGGACGCAATCGGAAGCTGTGGTATGCTGTGAGGTCGTAATCACTGCAATATTCG 1320
Qy     1321  TGTGCTCAAGCGCACTCCCGTCTCGGTAATGTTTTTGGCGCGACATCAATACGGTT 1380
Db     1321  TGTGCTCAAGCGCACTCCCGTCTCGGTAATGTTTTTGGCGCGACATCAATACGGTT 1380
Qy     1381  CTGGCAATATTCTGAATAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
Db     1381  CTGGCAATATTCTGAATAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
Qy     1441  ATTGTGAGCGGATAACAATTTACACAGAAACAGCCAGTCGCTTTAGGTGTTTTACGA 1500
Db     1441  ATTGTGAGCGGATAACAATTTACACAGAAACAGCCAGTCGCTTTAGGTGTTTTACGA 1500
Qy     1501  GCACCTTACCACCAAGGACCATAGATTATGAAAACTGAAGAAGGTAATCTGGTATCTGG 1560
Db     1501  GCACCTTACCACCAAGGACCATAGATTATGAAAACTGAAGAAGGTAATCTGGTATCTGG 1560
Qy     1561  ATTAACGGCGATAAAGGCTATACCGTCTCGCTGAGTCGGTAAGAAATTCAGAAAGAT 1620
Db     1561  ATTAACGGCGATAAAGGCTATACCGTCTCGCTGAGTCGGTAAGAAATTCAGAAAGAT 1620
Qy     1621  ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT 1680
Db     1621  ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT 1680
Qy     1681  GCGGCAATCGCGATGGCCCTGACATATCTTCTGGGCAACGACGCTTTGGTGCTAC 1740
Db     1681  GCGGCAATCGCGATGGCCCTGACATATCTTCTGGGCAACGACGCTTTGGTGCTAC 1740
Qy     1741  GCTCAATCTGGCGCTGTGGCTGAAATCAACCCGGCAAAAGCGTTCCAGGACAAGTGAT 1800
Db     1741  GCTCAATCTGGCGCTGTGGCTGAAATCAACCCGGCAAAAGCGTTCCAGGACAAGTGAT 1800
Qy     1801  CCGTTTACCTGGGATGCGGTACGTTTACAACGGCAAGCTGATTGCTTACCCGATCGTGT 1860
Db     1801  CCGTTTACCTGGGATGCGGTACGTTTACAACGGCAAGCTGATTGCTTACCCGATCGTGT 1860
Qy     1861  GAAGCGTTATCGCTGATTTTATAACAAGATCTGCTCGGAACCCGCCAAAAACCTGGGAA 1920
Db     1861  GAAGCGTTATCGCTGATTTTATAACAAGATCTGCTCGGAACCCGCCAAAAACCTGGGAA 1920
Qy     1921  GAGATCCCGCGCTGATTAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATTTCAAC 1980
Db     1921  GAGATCCCGCGCTGATTAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATTTCAAC 1980
```

```
Qy      1981  CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db      1981  CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Qy     2041  TATGAAGAACGCAAGTACGACATTAAGACGCTGGGCGTGGATAACGCTGGCGGAAACGG 2100
Db     2041  TATGAAGAACGCAAGTACGACATTAAGACGCTGGGCGTGGATAACGCTGGCGGAAACGG 2100
Qy     2101  GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAAACAACACATGAATGCAGACACGATTAC 2160
Db     2101  GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAAACAACACATGAATGCAGACACGATTAC 2160
Qy     2161  TCCATCGAGAAGCTGCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db     2161  TCCATCGAGAAGCTGCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy     2221  GCATGGTCCAAATCGACACACGCAAAAGTGAATTAATGGTGAACGCTACTGCGGACCTTC 2280
Db     2221  GCATGGTCCAAATCGACACACGCAAAAGTGAATTAATGGTGAACGCTACTGCGGACCTTC 2280
Qy     2281  AAGGTCACACCATCCAAACCGTTCTGTTGGCGTGTGAGCGAGGTAATTAACGCCCCAGT 2340
Db     2281  AAGGTCACACCATCCAAACCGTTCTGTTGGCGTGTGAGCGAGGTAATTAACGCCCCAGT 2340
Qy     2341  CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCGTG 2400
Db     2341  CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCGTG 2400
Qy     2401  GAAGCGGTTAATAAGACAAACCGCTGGTCCGTAGCGCTGAACTCTTACGAGGAAGAG 2460
Db     2401  GAAGCGGTTAATAAGACAAACCGCTGGTCCGTAGCGCTGAACTCTTACGAGGAAGAG 2460
Qy     2461  TTGCGAAGAGTCCACGCTATTGCGGCCCAACCATGAAAAACGCCAGAAAGGTGAATCATG 2520
Db     2461  TTGCGAAGAGTCCACGCTATTGCGGCCCAACCATGAAAAACGCCAGAAAGGTGAATCATG 2520
Qy     2521  CCGAACAATCCGCGAGATGTCGCTTTCTGGTATGCGCTGCGTACTGCGGTGATCAACGCC 2580
Db     2521  CCGAACAATCCGCGAGATGTCGCTTTCTGGTATGCGCTGCGTACTGCGGTGATCAACGCC 2580
Qy     2581  GCCAGCGTCTGACGCTGATGAAGCCCTGAAAGACGCCGACACTAAATTCGAGCTCG 2640
Db     2581  GCCAGCGTCTGACGCTGATGAAGCCCTGAAAGACGCCGACACTAAATTCGAGCTCG 2640
```

RESULT 15

AF031088 6706 bp DNA circular SYN 07-MAY-2001
Shuttle vector pMAL-pIII, complete sequence.

DEFINITION
AF031088

VERSION
AF031088.1 GI:2623821

KEYWORDS
Shuttle vector pMAL-pIII

SOURCE
Shuttle vector pMAL-pIII

ORGANISM
other sequences; artificial sequences; vectors.

REFERENCE
1 (bases 1 to 6706)

AUTHORS
Zwick,M.B., Bonnycastle,L.L., Noren,K.A., Venturini,S., Leong,E.,

Barbas,C.F. III, Noren,C.J. and Scott,J.K.

The maltose-binding protein as a scaffold for monovalent display of

peptides derived from phage libraries

Anal. Biochem. 264 (1), 87-97 (1998)

9902881

9784192

2 (bases 1 to 6706)

Noren,K.A. and Noren,C.J.

Direct Submission

Submitted (22-OCT-1997) New England Biolabs, 32 Tozer Road,

Beverly, MA 01915, USA

Location/Qualifiers

1..6706

source

/organism="Shuttle vector pMAL-pIII"

/mol_type="genomic DNA"

QY 1021 AGGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAAGAAAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAAGAAAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCTGGCAGCAGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCTGGCAGCAGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGAGTGCAATGCTTCTGGCG 1260
Db 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGAGTGCAATGCTTCTGGCG 1260
QY 1261 TCAGGCAGCATCGGAAGCTGGGTATGCTGTGAGGTGCTGAGGTGCTGATTAATTCG 1320
Db 1261 TCAGGCAGCATCGGAAGCTGGGTATGCTGTGAGGTGCTGAGGTGCTGATTAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACTCCCGTCTTGGATATGTTTTTGGCGGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTCTTGGATATGTTTTTGGCGGACATCATACGGTT 1380
QY 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGGA 1440
Db 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGGA 1440
QY 1441 ATTGTGAGGGATACAAATTTACACAGGAACGCGAGTCCGTTTAGTGTGTTTACGA 1500
Db 1441 ATTGTGAGGGATACAAATTTACACAGGAACGCGAGTCCGTTTAGTGTGTTTACGA 1500
QY 1501 GCACCTTCCACCAAGGACCATAG----- 1524
Db 1501 GCACCTTCCACCAAGGACCATAGCATATGAAAAATTAATTTCGCAATTCCTTTAGTG 1560
QY 1525 -----ATTATGAAAACTGGAAGAGTAAACTGTAATCTGG 1560
Db 1561 GTACCTTTCTATTCTCACTCGGCCGATATGAAATCGAAGAGGTAAACTGGTAATCTGG 1620
QY 1561 ATTAACGGGATAAAGGCTATAACGGTCTCGCTGAGTCCGTAAGAAATTCGAGAAAGT 1620
Db 1621 ATTAACGGGATAAAGGCTATAACGGTCTCGCTGAGTCCGTAAGAAATTCGAGAAAGT 1680
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACACAGTT 1680
Db 1681 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACACAGTT 1740
QY 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGCTAC 1740
Db 1741 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGCTAC 1800
QY 1741 GCTCAATCTGGCCCTGTTGGCTGAATCAACCCGGACAAAGCGTTCCAGACAAAGCTGTAT 1800
Db 1801 GCTCAATCTGGCCCTGTTGGCTGAATCAACCCGGACAAAGCGTTCCAGACAAAGCTGTAT 1860
QY 1801 CCGTTTACCTGGGATGCGGTAGCTTACAAAGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
Db 1861 CCGTTTACCTGGGATGCGGTAGCTTACAAAGCAAGCTGATTGCTTACCCGATCGCTGTT 1920
QY 1861 GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTCCGAAACCCGCAAAAAACCTGGGAA 1920
Db 1921 GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTCCGAAACCCGCAAAAAACCTGGGAA 1980
QY 1921 GAGATCCCGCGCTGATTAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1981 GAGATCCCGCGCTGATTAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 2040
QY 1981 CTGCAAGAACCGTACTTCACTGGCGGCTGATTGCTGCTGACGGGGTTTATCGGTTCAAG 2040
Db 2041 CTGCAAGAACCGTACTTCACTGGCGGCTGATTGCTGCTGACGGGGTTTATCGGTTCAAG 2100

QY 2041 TATGAAAAACGCAAGTAGCAGCATTTAAAGACGTGGCGCTGGATACCGTGGCGCGAAAGCG 2100
Db 2101 TATGAAAAACGCAAGTAGCAGCATTTAAAGACGTGGCGCTGGATACCGTGGCGCGAAAGCG 2160
QY 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2161 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAAACAAACACATGAATGCAGACACCGATTAC 2220
QY 2161 TCCATCGCAGAGCTGCTTTAATAAAGCGGCAACAGCGATGACCATCAACGGCGCGTGG 2220
Db 2221 TCCATCGCAGAGCTGCTTTAATAAAGCGGCAACAGCGATGACCATCAACGGCGCGTGG 2280
QY 2221 GCATGCTCCACATCGACACGAGCAAGTGAATTAATGGGTAAACGGTACTCCGACCTTC 2280
Db 2281 GCATGCTCCACATCGACACGAGCAAGTGAATTAATGGGTAAACGGTACTCCGACCTTC 2340
QY 2281 AAGGTCACACCATCCAAACCGTTGTTGGCGTGTGAGCGAGGTATTAAGCGCCCAAGT 2340
Db 2341 AAGGTCACACCATCCAAACCGTTGTTGGCGTGTGAGCGAGGTATTAAGCGCCCAAGT 2400
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
Db 2401 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2460
QY 2401 GAAGCGTTAATAAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTACGAGGAAAG 2460
Db 2461 GAAGCGTTAATAAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTACGAGGAAAG 2520
QY 2461 TTGGCGAAAGATCCACGTTATTGGCGCCACCATGGAACCGCCAGAAAGGTGAAATCATG 2520
Db 2521 TTGGCGAAAGATCCACGTTATTGGCGCCACCATGGAACCGCCAGAAAGGTGAAATCATG 2580
QY 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGCTGCTGCTGCTGATCAACGCC 2580
Db 2581 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGCTGCTGCTGATCAACGCC 2640
QY 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2641 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2700
QY 2641 AACACAAACAATAAACAATAAACAACCACTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2701 AACACAAACAATAAACAATAAACAACCACTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2760
QY 2701 GGATCCTCT 2709
Db 2761 GGATCCTCT 2769

Search completed: August 1, 2005, 17:21:49
Job time : 9285 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 6743.2 Seconds
(without alignments)
18627.964 Million cell updates/sec

Title: US-09-765-555B-14
Perfect score: 3300
Sequence: 1 ccgacacatcgatggtgc.....acgacgttcggactaagct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1:
2: gb_est2:
3: gb_hc:
4: gb_est3:
5: gb_est4:
6: gb_est5:
7: gb_est6:
8: gb_gss1:
9: gb_gss2:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	778	23.6	794	5	BQ751936
2	770.6	23.4	779	7	CK394523
3	768	23.3	769	7	CK118014
4	762	23.1	793	5	BQ751655
5	755.4	22.9	806	5	BQ751220
6	684.8	20.8	689	6	CB863541
7	676.4	20.5	712	1	AL038548
8	676	20.5	710	8	AY080106
9	647.4	19.6	650	1	AL044483
10	641	19.4	641	4	BJ063872
11	615.8	18.7	624	6	CB863814
12	612	18.5	794	1	AL045353
13	603	18.3	752	9	CL655575
14	589.8	17.9	608	8	AY080094
15	586.4	17.8	589	1	AL037742
16	581.2	17.6	601	1	AV594538
17	579	17.5	579	1	AL043868
18	570	17.3	782	7	CK394246
19	569	17.2	774	1	AL045337
20	563.2	17.1	585	9	CL658017
21	562.6	17.0	723	1	AL039416
22	549.4	16.6	753	1	AL045341
23	546.8	16.6	782	1	AL037051
24	545.4	16.5	547	6	CA895956

C 25	545	16.5	648	7	CK394349	CK394349	hgada2D0
C 26	540.2	16.4	569	1	AL039077	AL039077	DKF2p566G
C 27	534.4	16.2	644	6	CB862524	CB862524	HH04B08u
C 28	534	16.2	734	1	AL048427	AL048427	DKF2p586H
C 29	529	16.0	740	1	AL042909	AL042909	DKF2p434J
C 30	528	16.0	720	6	CA881974	CA881974	K0994C10-
C 31	523.2	15.9	528	1	AL038811	AL038811	DKF2p566O
C 32	519.4	15.7	521	6	CA890032	CA890032	B0136F03-
C 33	519	15.7	607	6	CB862099	CB862099	HH06A02Y
C 34	516.4	15.6	518	6	CB860082	CB860082	B0124F04-
C 35	510.2	15.5	686	1	AL044407	AL044407	DKF2p434D
C 36	509.4	15.4	511	6	CA895436	CA895436	B0192A07-
C 37	498	15.1	498	1	AL039076	AL039076	DKF2p566G
C 38	494.8	15.0	1025	1	AL038025	AL038025	DKF2p566C
C 39	487.6	14.8	498	7	CK394397	CK394397	hgada2H1
C 40	477.6	14.5	491	6	CB862715	CB862715	HH03F14u
C 41	465.8	14.1	679	1	AL039128	AL039128	DKF2p566K
C 42	462.8	14.0	490	1	AL039589	AL039589	DKF2p434D
C 43	458	13.9	458	6	CA887583	CA887583	B0137H04-
C 44	456.4	13.8	471	1	AL039649	AL039649	DKF2p434G
C 45	449.2	13.6	612	6	CB863409	CB863409	HH04K05W

ALIGNMENTS

RESULT 1
BQ751936
LOCUS BQ751936 794 bp mRNA linear EST 18-JUL-2002
DEFINITION EST632499 DSCT Colletotrichum trifolii cDNA clone pDST8-67, mRNA
sequence.
ACCESSION BQ751936
VERSION BQ751936.1 GI:21907341
KEYWORDS EST.
SOURCE Colletotrichum trifolii
ORGANISM Colletotrichum trifolii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
mitosporic Phyllachoraceae; Colletotrichum.
REFERENCE 1 (bases 1 to 794)
AUTHORS Samac,D.A., Dickman,M., Town,C.D., Van Aken,S., Utterback,T.,
Cheung,F. and Fraser,C.M.
TIGR sequence name: MTSN67TV More information is available at:
www.medicago.org
Seq primer: lgtA AtA Cga Ctc Act AtA ggg C).
Location/Qualifiers
1..794
/organism="Colletotrichum trifolii"
/mol_type="mRNA"
/strain="race 1"
/db_xref="taxon:5466"
/clone="pDST8-67"
/tissue_type="mycelia"
/dev_stage="Young, actively growing mycelia (3 days after
inoculation) grown in liquid culture (cutin minimal medium
containing 2%glucose)."
/lab_host="DH5alpha"
/clone_lib="DSTC"
/note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2:
EcoRI; isolate: 28p2; cDNA was prepared from polyA-
enriched RNA The cDNA was ligated into Lambda gptI from
Stratagene and packaged using Gigapack packaging extracts.
An aliquot of the amplified library was used to transduce

FEATURES
source
Contact: Deborah A. Samac
Department of Plant Pathology
University of Minnesota
495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
Tel: 612 625 1243
Fax: 651 649 5058
Email: debbys@puccini.crl.umn.edu
TIGR sequence name: MTSN67TV More information is available at:
www.medicago.org
Seq primer: lgtA AtA Cga Ctc Act AtA ggg C).
Location/Qualifiers
1..794
/organism="Colletotrichum trifolii"
/mol_type="mRNA"
/strain="race 1"
/db_xref="taxon:5466"
/clone="pDST8-67"
/tissue_type="mycelia"
/dev_stage="Young, actively growing mycelia (3 days after
inoculation) grown in liquid culture (cutin minimal medium
containing 2%glucose)."
/lab_host="DH5alpha"
/clone_lib="DSTC"
/note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2:
EcoRI; isolate: 28p2; cDNA was prepared from polyA-
enriched RNA The cDNA was ligated into Lambda gptI from
Stratagene and packaged using Gigapack packaging extracts.
An aliquot of the amplified library was used to transduce

E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN

Query Match 23.6%; Score 778; DB 5; Length 794;
Best Local Similarity 98.7%; Pred. No. 4.9e-219; Indels 0; Gaps 0;
Matches 784; Conservative 0; Mismatches 10;

Qy 3 GACACCATCGAATCGTGCAAAACCTTTCCGGTATGGCATATAGCGCCGGAAGAGAGT 62
Db 1 GACACTTTGGAATCGCGCAAAACCTTTCCGGTATGGCATATAGCGCCGGAAGAGAGT 60

Qy 63 CAATTCAGGTTGGTAATGTGAACACAGTAACGTTATACGATTCGACAGATATGCCGGT 122
Db 61 CAATTCAGGTTGGTAATGTGAACACAGTAACGTTATACGATTCGACAGATATGCCGGT 120

Qy 123 GTCTCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCAGTTTCTGCGAAAAACG 182
Db 121 GTCTCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCAGTTTCTGCGAAAAACG 180

Qy 183 CGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTACATTCCTCAACCGCGTGCGCAAA 242
Db 181 CGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTACATTCCTCAACCGCGTGCGCAAA 240

Qy 243 CAACCTCGCGGCAACAGCTGTTCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGCAC 302
Db 241 CAACCTCGCGGCAACAGCTGTTCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGCAC 300

Qy 303 GCGCGCTGCGAAATTTGTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCGTG 362
Db 301 GCGCGCTGCGAAATTTGTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCGTG 360

Qy 363 GTGGTGTGATGTGTAAGCAAGCGCGGTGAAAGCTGTAAGCGCGGTGCAAACTCTT 422
Db 361 GTGGTGTGATGTGTAAGCAAGCGCGGTGAAAGCTGTAAGCGCGGTGCAAACTCTT 420

Qy 423 CTCGCGCAACGCGTCAGTGGGCTGATCATTAACATTCGCTGGATGACAGGATGCCATT 482
Db 421 CTCGCGCAACGCGTCAGTGGGCTGATCATTAACATTCGCTGGATGACAGGATGCCATT 480

Qy 483 GCTGTGAAGCTGCTGCACATAATGTTCCGGCGTTATTTCTTGATGTTCTTGACAGACA 542
Db 481 GCTGTGAAGCTGCTGCACATAATGTTCCGGCGTTATTTCTTGATGTTCTTGACAGACA 540

Qy 543 CCCATCAACAGTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATCTG 602
Db 541 CCCATCAACAGTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATCTG 600

Qy 603 GTCGATTTGGTACACAGCAATCGCGTGTAGCGGCGCCATTAAAGTTCTGTCGCGCG 662
Db 601 GTCGATTTGGTACACAGCAATCGCGTGTAGCGGCGCCATTAAAGTTCTGTCGCGCG 660

Qy 663 CGTCTGGTCTGGCTGGCTGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
Db 661 CGTCTGGTCTGGCTGGCTGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 720

Qy 723 GAACGGGAAGCGACTGAGTGCCATGTCGGGTTTTCAACCAACCATGCAAAATGCTGAAT 782
Db 721 GAACGGGAAGCGACTGAGTGCCATGTCGGGTTTTCAACCAACCATGCAAAATGCTGAAT 780

Qy 783 GAGGGCATCGTTCC 796
Db 781 GAGGGCATCGTTCC 794

RESULT 2
CK394523/c
LOCUS
DEFINITION hggada4D08 Gland Cell Amplified cDNA Library Heterodera glycines
779 bp mRNA linear EST 30-DEC-2003

cDNA, mRNA sequence.
CK394523
CK394523.1 GI:40389794
EST.
Heterodera glycines
Heterodera glycines
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera.
1 (bases 1 to 779)
Wang,X., Allen,R., Gao,B., Goellner,M., Maier,T., Baum,T.,
Hussey,R. and Davis,E.
Submission: Wang, Allen, Gao, Goellner, Maier, Baum, Hussey, Davis
Unpublished (2003)
Contact: Tom Maier
Department of Plant Pathology, Baum Lab
Iowa State University
351 Bessey Hall, Ames, IA 50011, USA
Tel: 515-294-8854
Fax: 515-294-9420
Email: trmaier@iastate.edu
Heterodera glycines Gland Cell Amplified cDNA Library, single pass
sequence.

Location/Qualifiers
1..779
/organism="Heterodera glycines"
/mol_type="mRNA"
/db_xref="taxon:51029"
/cell_type="gland"
/dev_stage="mixed parasitic juvenile"
/clone_lib="Gland Cell Amplified cDNA Library"
/note="Organ: gland cell; Vector: pSportII"

ORIGIN

Query Match 23.4%; Score 770.6; DB 7; Length 779;
Best Local Similarity 99.4%; Pred. No. 7.6e-217; Indels 0; Gaps 0;
Matches 773; Conservative 0; Mismatches 5;

Qy 87 CCAGTAACGTTATACGATGTGCGAGAGTATGCGGTTCTCTTATCAGACCGTTTCCCGC 146
Db 779 CCAGTAACGTTATACGATGTGCGAGAGTATGCGGTTCTCTTATCAGACCGTTTCCCGC 720

Qy 147 GTGGTGAACAGCGCAGCCACGTTTTCGAAACCGCGGAAAAAGTGGAGCGCGGATG 206
Db 719 GTGGTGAACAGCGCAGCCACGTTTTCGAAACCGCGGAAAAAGTGGAGCGCGGATG 660

Qy 207 GCGGAGCTGAATTACATTTCCCAACCGCGTGGCAACAACCTGGCGGCAACAGTCGTTG 266
Db 659 GCGGAGCTGAATTACATTTCCCAACCGCGTGGCAACAACCTGGCGGCAACAGTCGTTG 600

Qy 267 CTGATTGGCGTTGCGACCTCCAGTCTGGCCCTGCACGCGCCGTCGCAAAATTTGTCGCGCG 326
Db 599 CTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCACGCGCCGTCGCAAAATTTGTCGCGCG 540

Qy 327 ATTAATCTCGCGCGATCAACTGGGTGCGAGGTGGTGTGTCGATGGTAGAAGCAAGC 386
Db 539 ATTAATCTCGCGCGATCAACTGGGTGCGAGGTGGTGTGTCGATGGTAGAAGCAAGC 480

Qy 387 GCGTTCGAAGCTGTAAGCGGCGTGCACAATCTTCTCGCAACCGCGTCAGTGGGCTG 446
Db 479 GCGTTCGAAGCTGTAAGCGGCGTGCACAATCTTCTCGCAACCGCGTCAGTGGGCTG 420

Qy 447 ATCATTAACTATCCGCTGGATGACAGGATGCCATTTGCTGGAAGCTGCTGCACATAA 506
Db 419 ATCATTAACTATCCGCTGGATGACAGGATGCCATTTGCTGGAAGCTGCTGCACATAA 360

Qy 507 GTTCCGCGTTATTTCTTGATGTCTGACAGACACCCATCAACAGTATATTTCTCC 566
Db 359 GTTCCGCGTTATTTCTTGATGTCTGACAGACACCCATCAACAGTATATTTCTCC 300

Qy 567 CATGAAGACGGTACGCGACTGGCGTGGAGCATCTGGTCGATTGGGTCAACAGCAATC 626
Db 299 CATGAAGACGGTACGCGACTGGCGTGGAGCATCTGGTCGATTGGGTCAACAGCAATC 240

COMMENT

Other ESTs: EST632217
 Contact: Deborah A. Samac
 Department of Plant Pathology
 University of Minnesota
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
 Tel: 612 625 1243
 Fax: 651 649 5058
 Email: debbys@puccini.crl.umn.edu
 TIGR sequence name: MTSAF59TV More information is available at:
 www.medicago.org
 Seq primer: (gta Ata Cga Ctc Act Ata ggg C).

FEATURES

Location/Qualifiers
 1..793
 /organism="Colletotrichum trifolii"
 /mol_type="mRNA"
 /strain="race 1"
 /db_xref="taxon:5466"
 /clone="pDSC76-59"
 /tissue_type="mycelia"
 /dev_stage="young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
 /lab_host="DH5alpha"
 /clone_lib="DSC7"
 /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2sp2 ; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gtl1 from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN

Query Match
 Best Local Similarity 23.1%; Score 762; DB 5; Length 793;
 Matches 777; Conservative 0; Mismatches 5; Indels 2; Gaps 1;
 3 GACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCCCGGAAGAGAGT 62
 1 GACATCTTCGATGGCGCAAAACCTTTCGCGGTATGGCATGATAGCCCGGAAGAGAGT 60
 63 CAATTCAGGGTGGTAATGTGAACACCGATTAACGATGTCGAGAGTATGCCCGT 122
 61 CAATTCAGGGTGGTAATGTGAACACCGATTAACGATGTCGAGAGTATGCCCGT 120
 123 GTCTCTTATCAGACCGTTTCGCGGTGGTGAACCGCCGACGCTTCTGCGAANAACG 182
 121 GTCTCTTATCAGACCGTTTCGCGGTGGTGAACCGCCGACGCTTCTGCGAANAACG 180
 183 CGGGAANAAGTGAAGCGCGCATGGCGAGTCGAATTCACATTCACACCGCGTGGCACAA 242
 181 CGGGAANAAGTGAAGCGCGCATGGCGAGTCGAATTCACATTCACACCGCGTGGCACAA 240
 243 CAACCTGGCGGCAACAGTGTGCTGATGGCGGTGGCCACCTCCAGTTCGGCCCTGCAC 302
 241 CAACCTGGCGGCAACAGTGTGCTGATGGCGGTGGCCACCTCCAGTTCGGCCCTGCAC 300
 303 CGCGCGTCGCAAAATGTTCGCGGCGCATTAATTCGCGCGCATCACTGGGTGGCCAGCTG 362
 301 CGCGCGTCGCAAAATGTTCGCGGCGCATTAATTCGCGCGCATCACTGGGTGGCCAGCTG 360
 363 GTGGTGTGATGGTAGAACGAGCGGCTCGAAGCGCTTAAAGCGCGGTGACACATCTT 422
 361 GTGGTGTGATGGTAGAACGAGCGGCTCGAAGCGCTTAAAGCGCGGTGACACATCTT 420
 423 CTCGCGCAACCGCTCAGTGGGCTGATCACTTAATCTCCGTGGATGACGAGATGCCATT 482
 421 CTCGCGCAACCGCTCAGTGGGCTGATCACTTAATCTCCGTGGATGACGAGATGCCATT 480

QY 483 GCTGTGGAAGCTGCTGCACATAATGTTCCGCGTTATTTCTTGATGCTCTGACACAGACA 542
 Db 481 GCTGTGGAAGCTGCTGCACATAATGTTCCGCGTTATTTCTTGATGCTCTGACACAGACA 540
 QY 543 CCCATCAACAGATATATTTCTCCCATGAAGACGGTACGGGCTGGGAGCATCTG 602
 Db 541 CCCATCAACAGATATATTTCTCCCATGAAGACGGTACGGGCTGGGAGCATCTG 600
 QY 603 GTCCATTGGGTACACCAAAATCGCGCTGTAGCGGCGCCATTAATTTCTGTCGCGG 662
 Db 601 GTCCATTGGGTACACCAAAATCGCGCTGTAGCGGCGCCATTAATTTCTGTCGCGG 660
 QY 663 CGTCTGCGTCTGCGTGGCTGGCATATAATCTCACTCGCAATCAAAATTCAGCGATAGCG 722
 Db 661 CGTCTGCGTCTGCGTGGCTGGCATATAATCTCACTCGCAATCAAAATTCAGCGATAGCG 720
 QY 723 GAACGGGAAGCGGACCTGAGTGCATGTCGGTTCACAAACCAACCATGCAATGCTGAAT 782
 Db 721 GAACGGGAAGCGGACCTGAGTGCATGTCGGTTCACAAACCAACCATGCAATGCTGAAT 778
 QY 783 GAGG 786
 Db 779 GAGG 782

RESULT 5
 LOCUS BQ751220
 DEFINITION BQ751220 DSC7 Colletotrichum trifolii cDNA clone pDSC73-64, mRNA sequence.
 ACCESSION BQ751220
 VERSION BQ751220.1 GI:21906625
 KEYWORDS EST.
 SOURCE Colletotrichum trifolii
 ORGANISM Colletotrichum trifolii
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae; mitosporic Phyllachoraceae; Colletotrichum.
 1 (bases 1 to 806)
 Samac, D.A., Dickman, M., Town, C.D., Van Aken, S., Utterback, T., Cheung, F. and Fraser, C.M.
 ESTs from mycelia of Colletotrichum trifolii race 1
 Unpublished (2002)
 Other ESTs: EST631782
 Contact: Deborah A. Samac
 Department of Plant Pathology
 University of Minnesota
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
 Tel: 612 625 1243
 Fax: 651 649 5058
 Email: debbys@puccini.crl.umn.edu
 TIGR sequence name: MTSAC64TV More information is available at:
 www.medicago.org
 Seq primer: (gta Ata Cga Ctc Act Ata ggg C).

FEATURES

Location/Qualifiers
 1..806
 /organism="Colletotrichum trifolii"
 /mol_type="mRNA"
 /strain="race 1"
 /db_xref="taxon:5466"
 /clone="pDSC73-64"
 /tissue_type="mycelia"
 /dev_stage="young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
 /lab_host="DH5alpha"
 /clone_lib="DSC7"
 /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2sp2 ; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gtl1 from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid


```

QY 441 GGGCTGATCAATTAATATCCGCTGGATGACACAGGATGCCATTGCTGTGGAGCTGCCTGC 500
Db |||||||
QY 269 GGGCTGATCAATTAATATCCGCTGGATGACACAGGATGCCATTGCTGTGGAGCTGCCTGC 210
Db |||||||
QY 501 ACTAATGTTCCGGCGTTATTTCTTATGATCTCTGACACAGACACCCCATCAACAGTATTATT 560
Db |||||||
QY 209 ACTAATGTTCCGGCGTTATTTCTTATGATCTCTGACACAGACACCCCATCAACAGTATTATT 150
QY 561 TTCTCCCATGAAGACGGTACGGACTCGGCGTGGAGCATCTGGTCGCAATGGGTCACACAG 620
Db |||||||
QY 149 TTCTCCCATGAAGACGGTACGGACTCGGCGTGGAGCATCTGGTCGCAATGGGTCACACAG 90
QY 621 CAATCGCGCTGTTAGCGGCGCCATTAAGTTCTGTCTCGGCGCGTCTCGCTGCTGCTGCGC 680
Db |||||||
QY 89 CAATCGCGCTGTTAGCGGCGCCATTAAGTTCTGTCTCGGCGCGTCTCGCTGCTGCTGCTGCGC 30
QY 681 TGGCATTAATATCTCACTCGCAATCAAT 709
Db |||||||
QY 29 TGGCATTAATATCTCACTCGCAATCAAT 1

RESULT 7
LOCUS AL038548/c 712 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZps566E0346_r1 566 (synonym: hfkd2) Homo sapiens cDNA clone
AL038548
ACCESSION AL038548.1 GI:5407738
VERSION DKFZps566E0346.5, mRNA sequence.
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 712)
JOURNAL Ottenwaelder, B., Obermaier, B., Mewes, H.W., Gassenhuber, J. and
COMMENT EST (Ottenwaelder, et al.)
Unpublished (1999)
Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by MediGenomix (Martinried/Germany) within the cDNA
sequencing consortium of the German Genome Project. No s1 sequence
available.
This clone (DKFZps566E0346) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin- Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source
Location/Qualifiers
1..712
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZps566E0346"
/tissue_type="kidney"
/dev_stage="fetal"
/lab_host="X1-2blue"
/clone_lib="566 (synonym: hfkd2)"
/note="vector: pAMPI; Site_1: NotI; Site_2: SalI"

ORIGIN
Query Match 20.5%; Score 676.4; DB 1; Length 712;
Best Local Similarity 99.6%; Pred. No. 7.2e-189;
Matches 699; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 33 GGTATGCGATGATAGCCCGGAGAGAGTCATTCAGGGTGGTGATGTGAACACAGTA 92
Db |||||||
QY 712 GGTATGCGATGATAGCCCGGAGAGAGTCATTCAGGGTGGTGATGTGAACACAGTA 653
QY 93 ACGTTATACATGTCGAGAGTATGCCGGTGTCTCTTATCAGACCGGTTCCCGCGTGTG 152
Db |||||||

```

```

Db 652 ACGTTATACATGTCGACAGATATGCCGGTGTCTCTTATCAGACCGTTTCCCGCGTGTG 593
QY 153 AACAGCGCAGCCAGTTTCTGCGAAACCGCGG-AAAAAGTGAAGCGCGATGGCGGA 211
Db |||||||
QY 592 AACAGCGCAGCCAGTTTCTGCGAAACCGCGGAAAAAGTGAAGCGCGATGGCGGA 533
QY 212 GCTCAATTAACATTCCTCAACCGCGTGGCACAACACTGGCGGCAACAGTGTGCTGAT 271
Db |||||||
QY 532 GCTCAATTAACATTCCTCAACCGCGTGGCACAACACTGGCGGCAACAGTGTGCTGAT 473
QY 272 TGGCGTTGCCACCTCCAGTCTTGGCCCTTCGCAC-GGCGCGTGCCTGCTGCGCGGATTA 330
Db |||||||
QY 472 TGGCGTTGCCACCTCCAGTCTTGGCCCTTCGCACGGCGCGTGCCTGCGCGGATTA 413
QY 331 AATCTCGCGCGATCAACTGGGTGCCAGCGTGGTGTGTCGATGGTGAACGAAAGCGCG 390
Db |||||||
QY 412 AATCTCGCGCGATCAACTGGGTGCCAGCGTGGTGTGTCGATGGTGAACGAAAGCGCG 353
QY 391 TCGAAGCGCTGTAAAGCGCGGTGCACAATCTTCTCGCGCAACCGGTGAGTGGGCTGATCA 450
Db |||||||
QY 352 TCGAAGCGCTGTAAAGCGCGGTGCACAATCTTCTCGCGCAACCGGTGAGTGGGCTGATCA 293
QY 451 TTAACATATCCGCTGGATGACACAGGATGCCATTGCTGTGGAAGCTGCTGCACCTAAATGTC 510
Db |||||||
QY 292 TTAACATATCCGCTGGATGACACAGGATGCCATTGCTGTGGAAGCTGCTGCACCTAAATGTC 233
QY 511 CGGCGTTATTTCTGATGTCCTGACAGACACCCCATCAACAGTATTATTTCTCCCATG 570
Db |||||||
QY 232 CGGCGTTATTTCTGATGTCCTGACAGACACCCCATCAACAGTATTATTTCTCCCATG 173
QY 571 AAGACGGTACCGCATGGCGGTGGAGCATCTGTCGCAATGGGTGCAACGCAATCCGCG 630
Db |||||||
QY 172 AAGACGGTACCGCATGGCGGTGGAGCATCTGTCGATTTGGGTGTCACGCAATCCGCG 113
QY 631 TGTAGCGGCGCCATTAAGTTCTGTCGCGCGGTCTGCTGCTGGTGGTGGCATAAAT 690
Db |||||||
QY 112 TGTAGCGGCGCCATTAAGTTCTGTCGCGCGGTCTGCTGCTGGTGGTGGCATAAAT 53
QY 691 ATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAG 732
Db |||||||
QY 52 ATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAG 11

RESULT 8
LOCUS AY080106 710 bp DNA linear GSS 06-NOV-2002
DEFINITION AY080106 Scripps Pier (La Jolla, CA) uncultured virus community
uncultured marine virus genomic clone SI05lp36L, genomic survey
sequence.
ACCESSION AY080106
VERSION AY080106.1 GI:24745302
KEYWORDS GSS.
SOURCE uncultured marine virus
ORGANISM uncultured marine virus
REFERENCE 1 (bases 1 to 710)
AUTHORS Viruses; environmental samples.
Breitbart, M., Salamon, P., Andresen, B., Mahaffy, J.M., Segall, A.M.,
Mead, D., Azam, F. and Rohwer, F.
Genomic analysis of uncultured marine viral communities
Proc. Natl. Acad. Sci. U.S.A. 99 (22), 14250-14255 (2002)
JOURNAL 22294988
MEDLINE 12384570
PUBMED 12384570
COMMENT Contact: Rohwer F
Biology Dept.
San Diego State University
5500 Campanile Dr, San Diego, CA 92102, USA
Tel: 6195941336
Fax: 619595676
Email: forest@sunstroke.sdsu.edu
Class: shotgun.
Location/Qualifiers
1..710
/organism="uncultured marine virus"

FEATURES
source

```

/mol_type="genomic DNA"
/db_xref="taxon:186617"
/clone="SI051pJG6L"
/clone_lib="Scripps Pier (La Jolla, CA) uncultured virus community"
/note="Marine viruses were isolated from 200 liters of surface seawater using a combination of differential filtration and density-dependent gradient centrifugation. Linker-amplified shotgun libraries were constructed by randomly shearing the total marine viral community DNA, end-repairing, ligating dsDNA linkers to the ends, and amplifying the fragments using Vent DNA polymerase. The resulting fragments were ligated into the pSMART vector and electroporated into MC12 cells (Lucigen; Middleton, WI)."

ORIGIN

Query Match 20.5%; Score 676; DB 8; Length 710;
Best Local Similarity 98.9%; Pred. No. 9.5e-189;
Matches 690; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
Qy 400 GTAAAGCGCGGTGACAAATCTTCTCGGCAACGGGTGAGTGGCTGATCAATTAATCACTATC 459
Db |||||
8 GTGCTGACGCGTGCACAACTCTTCTCGGCAACGGGTGAGTGGCTGATCAATTAATCACTATC 67
Qy 460 CGCTGGATGACAGGATGCTATGCTGTGGAAAGCTGCTGCACTAATGTTCCGGCGTTAT 519
Db |||||
68 CGCTGGATGACAGGATGCTATGCTGTGGAAAGCTGCTGCACTAATGTTCCGGCGTTAT 127
Qy 520 TTCTTGATGCTCTGACAGACACCCATCAACAGTATTATTTCTCCCATGAAGACGGTA 579
Db |||||
128 TTCTTGATGCTCTGACAGACACCCATCAACAGTATTATTTCTCCCATGAAGACGGTA 187
Qy 580 CGCGACTGGCGTGGAGCATCTGTGCGATTTGGGTCAACAGCAATCGCGCTGTTAGCGG 639
Db |||||
188 CGCGACTGGCGTGGAGCATCTGTGCGATTTGGGTCAACAGCAATCGCGCTGTTAGCGG 247
Qy 640 GCCATTAAGTTCTGTCTCGGGCGCTGCTGCTGCTGGCTGGCTGGCAATAAATCACTC 699
Db |||||
248 GCCATTAAGTTCTGTCTCGGGCGCTGCTGCTGCTGGCTGGCTGGCAATAAATCACTC 307
Qy 700 GCAATCAATTCAGCCGATAGCGGAACGGGAGCGAGTGGAGTCCCATGTCGGGTTTC 759
Db |||||
308 GCAATCAATTCAGCCGATAGCGGAACGGGAGCGAGTGGAGTCCCATGTCGGGTTTC 367
Qy 760 AACAAACCATGCAATGCTGAATGAGGGCATCGTTCCCACTCGCATGCTGTTGTCGAACG 819
Db |||||
368 AACAAACCATGCAATGCTGAATGAGGGCATCGTTCCCACTCGCATGCTGTTGTCGAACG 427
Qy 820 ATCAGATGGCGTGGCGCAATGCGCGCAATTACCGAGTTCGGGCTGCGCGTGGTGGCGG 879
Db |||||
428 ATCAGATGGCGTGGCGCAATGCGCGCAATTACCGAGTTCGGGCTGCGCGTGGTGGCGG 487
Qy 880 ATATCTCGGTAGTGAGTACGACGATACCGAAGACAGTCAATGTTATATCCCGCGTTAA 939
Db |||||
488 ATATCTCGGTAGTGAGTACGACGATACCGAAGACAGTCAATGTTATATCCCGCGTTAA 547
Qy 940 CCACCATCAACAGGATTTTCGCTGCTGGGCAACACAGCGTGGACCGTTCGTGCAAC 999
Db |||||
548 CCACCATCAACAGGATTTTCGCTGCTGGGCAACACAGCGTGGACCGTTCGTGCAAC 607
Qy 1000 TCTCTCAGGGCCAGCGGTGAAGGGCAATFACGTGTGTCGGCTCTCACTGGTGAAGAA 1059
Db |||||
608 TCTCTCAGGGCCAGCGGTGAAGGGCAATFACGTGTGTCGGCTCTCACTGGTGAAGAA-AA 666
Qy 1060 AAACACCTGCGGCCCAATACGCAACCGCCTCTCCC 1097
Db |||||
667 AAACACCTGCGGCCCAATACGCAACCGCCTCTCCC 704

RESULT 9
AL044483/c
LOCUS

AL044483 660 bp mRNA linear EST 04-SEP-2003

DEFINITION DKFZp434i1102_s1 434 (synonym: htes3) Homo sapiens cDNA clone
DKFZp434i1102 3', mRNA sequence.
ACCESSION AL044483
VERSION AL044483.1 GI:5432701
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 660)
AUTHORS Ansoorge, W., Benes, V., Krieger, S., Mewes, H.W., Gassenhuber, J. and Wismann, S.
TITLE EST (Ansoorge, Benes, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 3' sequence of the clone insert
Clone from S. Wismann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wismann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No r1 sequence available.
This clone (DKFZp434i1102) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
Location/Qualifiers
1..660
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp434i1102"
/tissue type="testis"
/dev stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 19.6%; Score 647.4; DB 1; Length 660;
Best Local Similarity 99.7%; Pred. No. 2.9e-180;
Matches 659; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 85 AACCAAGTAACTTATACGATGTCGAGAGTATGCGGTGTCTCTTATCAGACCGTTTCCC 144
Db |||||
660 AACCAAGTAACTTATACGATGTCGAGAGTATGCGGTGTCTCTTATCAGACCGTTTCCC 601
Qy 145 CGTGTGTAACAGCGCCAGCCACGTTTCTGCGAAACCGGGGAAAAAGTGGAGCGGGA 204
Db |||||
600 CGTGTGTAACAGCGCCAGCCACGTTTCTGCGAAACCGGGGAAAAAGTGGAGCGGGA 541
Qy 205 TGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAACAACCTGCGGGGCAACAGTGT 264
Db |||||
540 TGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAACAACCTGCGGGGCAACAGTGT 481
Qy 265 TGCTGATTGGGTGGCCACCTCCAGTCTGGCCCTGCAACGCGCGTTCGCAAAATGTCGCGG 324
Db |||||
480 TGCTGATTGGGTGGCCACCTCCAGTCTGGCCCTGCAACGCGCGTTCGCAAAATGTCGCGG 421
Qy 325 CGATTAAATCTCGCGCCCGATCAACTGGGTGCGAGGTGGTGTGCGATGGTAGAACGAA 384
Db |||||
420 CGATTAAATCTCGCGCCCGATCAACTGGGTGCGAGGTGGTGTGCGATGGTAGAACGAA 361
Qy 385 CGCGGTGCAAGCGCTGTAAGCGGGTGCAACATCTTCTCGCGCAACGCGTCAGTGGGC 444
Db |||||
360 CGCGGTGCAAGCGCTGTAAGCGGGTGCAACATCTTCTCGCGCAACGCGTCAGTGGGC 301
Qy 445 TGATCATTTAACTATCCGCTGGATGACACAGGATGCCATTGCTGTGGAAGCTGCTGCACCTA 504
Db |||||
300 TGATCATTTAACTATCCGCTGGATGACACAGGATGCCATTGCTGTGGAAGCTGCTGCACCTA 241
Qy 505 ATGTTCCGGCGTTATTTCTTGATGTCTCTGACGACACCCATCAACAGTATTATTTTCT 564

```

Db      240 ATGTTCCGGCGTTATTTCTTGATGTTCTTGACAGACCCATCAACAGTATTTTCT 181
QY      565 CCATGAAGACGGTAGCGACTGGCGGTGGAGCATCTGTCATTTGGGTACACAGCAAA 624
Db      180 CCATGAAGACGGTAGCGACTGGCGGTGGAGCATCTGTCGCAATTTGGGTACACAGCAA 121
QY      625 TCGCGTGTAGCGGCCCAATTAAGTTCTGTCGCGCGTCTGCGTCTGGGTGCTGGCTGCG 684
Db      120 TCGCGTGTAGCGGCCCAATTAAGTTCTGTCGCG- GGTCTGCGTCTGGGTGCTGGCTGCG 62
QY      685 ATAAATATCTACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGCCGACTGGAGTG 744
Db      61 ATAAATATCTACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGCCGACTGGAGTG 2
QY      745 C 745
Db      1 C 1

RESULT 10
BJ063872/c
LOCUS
DEFINITION
BJ063872 NIBB Mochii normalized Xenopus tailbud library Xenopus
laevis cDNA clone XL077K02 5', mRNA sequence.
ACCESSION
BJ063872
VERSION
BJ063872.1 GI:17471074
KEYWORDS
SOURCE
ORGANISM
Xenopus laevis (African clawed frog)
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Xenopus.
1 (bases 1 to 641)
Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and
Kohara,Y.
Expressed genes in X. laevis embryo
Unpublished (2001)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp
The information of this clone is available through the following
URL.
http://xenopus.nibb.ac.jp.
FEATURES
source
Location/Qualifiers
1..641
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="XL077K02"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/clone_lib="NIBB Mochii normalized Xenopus tailbud
library"
ORIGIN
Query Match 19.4%; Score 641; DB 4; Length 641;
Best Local Similarity 100.0%; Pred.No. 2.3e-178;
Matches 641; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 30 CGCGGTATGCATGATAGCCCGGAGAGAGATCAATTCAGGTGGTGAATGTGAACCA 89
Db 641 CGCGGTATGCATGATAGCCCGGAGAGAGATCAATTCAGGTGGTGAATGTGAACCA 582
QY 90 GTAACGTTATACGATGTCGAGAGATAGCCGGTGTCTTTATCAGACCGTTTCCCGGTG 149
Db 581 GTAACGTTATACGATGTCGAGAGATAGCCGGTGTCTTTATCAGACCGTTTCCCGGTG 522
QY 150 GTGAACACCGCCAGCCACGTTTCTCGCAAAACCGGGGAAAAAGTGAAGCGCGATGGCG 209

```

```

Db      521 GTGAACAGGCGCAGCCACGTTTCTCGCAAAACCGGGGAAAAAGTGAAGCGCGATGGCG 462
QY      210 GAGCTGAATTAACATTCCTCAACCGGTGGCAACAACTGGCGGGCAACAGTTCGTTGCTG 269
Db      461 GAGCTGAATTAACATTCCTCAACCGGTGGCAACAACTGGCGGGCAACAGTTCGTTGCTG 402
QY      270 ATTGGCGTTGCACTTCAAGTCTGCGCCTGCAAGCGCGCTCGCAAAATTTGTCGGCGGAT 329
Db      401 ATTGGCGTTGCACTTCAAGTCTGCGCCTGCAAGCGCGCTCGCAAAATTTGTCGGCGGAT 342
QY      330 AAATCTCGCGCGGATCAACTGGGTGGCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 389
Db      341 AAATCTCGCGCGGATCAACTGGGTGGCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 282
QY      390 GTGAAGCCTGTAAAGCGCGGTGCACAATCTTCTCGCGCAAGCGCTCAGTGGGCTGATC 449
Db      281 GTGAAGCCTGTAAAGCGCGGTGCACAATCTTCTCGCGCAAGCGCTCAGTGGGCTGATC 222
QY      450 ATTAACCTATCGCTGGATGACAGGATGCCATTCCTGTTGGAAGTGCCTGCATTAATGTT 509
Db      221 ATTAACCTATCGCTGGATGACAGGATGCCATTCCTGTTGGAAGTGCCTGCATTAATGTT 162
QY      510 CCGCGGTTATTTCTTGATGTTCTTGACGACACACCATCAACAGTATTTTCTCCCAT 569
Db      161 CCGCGGTTATTTCTTGATGTTCTTGACGACACACCATCAACAGTATTTTCTCCCAT 102
QY      570 GAAGACGGTACGGACTGGCGGTGGAGCATCTGTCGCAATTTGGGTCAACAGCAATTCGCG 629
Db      101 GAAGACGGTACGGACTGGCGGTGGAGCATCTGTCGCAATTTGGGTCAACAGCAATTCGCG 42
QY      630 CTGTTAGCGGGCCATTAGTCTGTTGTTGGCGCGTCTGGG 670
Db      41 CTGTTAGCGGGCCATTAGTCTGTTGTTGGCGCGTCTGGG 1

RESULT 11
CB863814/c
LOCUS
DEFINITION
CB863814 HH Hordeum vulgare cDNA clone HH07B20 3-PRIME, mRNA
sequence.
ACCESSION
CB863814
VERSION
CB863814.1 GI:30058373
KEYWORDS
SOURCE
ORGANISM
Hordeum vulgare
Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
1 (bases 1 to 624)
Varshney,R.K., Zhang,H., Burton,R., Stein,N., Langridge,P. and
Graner,A.
Barley ESTs from coleoptile tissue
Unpublished (2003)
Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 624 Std Error: 0.00
Plate: 7 row: B column: 20
Seq primer: SP6.
Location/Qualifiers
1..624
/organism="Hordeum vulgare"
/mol_type="mRNA"
/cultivar="Sloop"
/db_xref="GABI:555471"
/db_xref="taxon:4513"
/clone="HH07B20"
/tissue_type="coleoptile"
FEATURES
source

```

/dev stage="coleoptile, 1 day old"	
/lab_host="DH10B"	
/clone_lib="HH"	
/note="Vector: pSPORT; Site 1: SalI (5-end of cDNA); Site 2: NotI (3-end of cDNA); Due to the cloning system used 'blue/white selection for recombinants is not 100 % reliable. Average insert size is 1.3 kb."	
ORIGIN	
Query Match	18.7%; Score 615.8; DB 6; Length 624;
Best Local Similarity	99.7%; Pred. NO. 7e-171;
Matches 617; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	3 GACACATCGAATGCTGCAGAAACCTTTTCGGGTATGGCATGATAGCGCCGCGAAGAGAGT 62
Db	
619	GACACATCGAATGCTGCAGAAACCTTTTCGGGTATGGCATGATAGCGCCGCGAAGAGAGT 560
QY	63 CAATTACAGGTTGCTCAATGTGAACACAGTAACGTTATACGATGTCGACAGATATCCCGGT 122
Db	
559	CAATTACAGGTTGCTCAATGTGAACACAGTAACGTTATACGATGTCGACAGATATCCCGGT 500
QY	123 GTCTCTTATCAGACCGTTTCCTCGCGTGTGAACACGAGCCAGCCAGCTTTTCTCGGAAACG 182
Db	
499	GTCTCTTATCAGACCGTTTCCTCGCGTGTGAACACGAGCCAGCCAGCTTTTCTCGGAAACG 440
QY	183 CGGGAAAAGTGGAAAGCGCGGATGCGGAGCTGAATTACATTTCCCAACCGGTGGCAAA 242
Db	
439	CGGGAAAAGTGGAAAGCGCGGATGCGGAGCTGAATTACATTTCCCAACCGGTGGCAAA 380
QY	243 CAACTGGCGGCAACACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCAC 302
Db	
379	CAACTGGCGGCAACACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCAC 320
QY	303 GCGCGTGCCTAAATGTCGCGCGATTAAATCTCGCGCGGATCAACTGGGTGCCAGCGTG 362
Db	
319	GCGCGTGCCTAAATGTCGCGCGATTAAATCTCGCGCGGATCAACTGGGTGCCAGCGTG 260
QY	363 GTGGTGTGATGTTAGAACGAGCGCGTGAAGCCTGTAAAGCGCGGTGCACAACTTT 422
Db	
259	GTGGTGTGATGTTAGAACGAGCGCGTGAAGCCTGTAAAGCGCGGTGCACAACTTT 200
QY	423 CTCGGCAACCGTCAGTGGGCTGATCAATTAATCTCGCTGGATGCCAGATGCCATT 482
Db	
199	CTCGGCAACCGTCAGTGGGCTGATCAATTAATCTCGCTGGATGCCAGATGCCATT 140
QY	483 GCTGGGAAGCTGCTGCACCTAATGTTCCGGGTTATTTCTTGATGTCTTGACAGACA 542
Db	
139	GCTGGGAAGCTGCTGCACCTAATGTTCCGGGTTATTTCTTGATGTCTTGACAGACA 80
QY	543 CCCATCAACAGATTTATTTTCCCATGAAGACGTAACGCTGCGGCTGGAGCATCTG 602
Db	
79	CCCATCAACAGATTTATTTTCCCATGAAGACGTAACGCTGCGGCTGGAGCATCTG 20
QY	603 GTCGCAATGGGTCAACAGC 621
Db	
19	GTGCAATGGGTCAACAGC 1
RESULT 12	
AL045353/c	
LOCUS	
DEFINITION	
DKEP434B075_e1 434 (synonym: htes3) Homo sapiens cDNA clone	
DKEP434B075 3', mRNA sequence.	
ACCESSION	
AL045353	
VERSION	
AL045353.1 GI:5433506	
KEYWORDS	
EST.	
SOURCE	
Homo sapiens (human)	
ORGANISM	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	
1 (bases 1 to 794)	
AUTHORS	
Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann,S.	
TITLE	
JOURNAL	
COMMENT	
EST (Duesterhoeft, et al.)	
Unpublished (1999)	
Contact: MIPS	
MIPS	
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany	
This is the 3' sequence of the clone insert	
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); sequenced by Qiagen within the cDNA sequencing consortium of the German Genome Project. rl sequence also available.	
This clone is available at the RZPD in Berlin.	
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.	
FEATURES	
source	
1..794	
/organism="Homo sapiens"	
/mol_type="mRNA"	
/db_xref="taxon:9606"	
/clone="DKEP434B075"	
/tissue_type="testis"	
/dev_stage="adult"	
/lab_host="DH10B"	
/clone_lib="434 (synonym: htes3)"	
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI"	
ORIGIN	
Query Match	18.5%; Score 612; DB 1; Length 794;
Best Local Similarity	99.7%; Pred. NO. 1e-169;
Matches 634; Conservative	0; Mismatches 0; Indels 2; Gaps 2;
QY	569 TGAAGACGGTACGCGACTGGCGGTGGAGCATCTGGTCGTCATGGTCCACGCAATCGC 628
Db	
794	TGAAGACGGTACGCGACTGGCGGTGGAG-ATCTGGTCGATGGGTCCACGCAATCGC 736
QY	629 GCTGTTAGCGGCGCCATTAAAGTTCTCTCGCGCGCTCTGCGTCTGGCTGGCTGGCATAA 688
Db	
735	GCTGTTAGCGGCGCCATTAAAGTTCTGCTCGCGGCTCTG-GTCTGGCTGGCTGGCATAA 677
QY	689 ATATCTCACTCGCAATCAAATTCAGCCGATAGCGAACGGGAAGCGCATGGAGTGCAT 748
Db	
676	ATATCTCACTCGCAATCAAATTCAGCCGATAGCGAACGGGAAGCGCATGGAGTGCAT 617
QY	749 GTCCGGTTTTCAACAAACCATGCAATGCTGAATGAGGGCATCGTCCACATGCCATGCT 808
Db	
616	GTCCGGTTTTCAACAAACCATGCAATGCTGAATGAGGGCATCGTTCACATGCCATGCT 557
QY	809 GGTTCGCCAACGATCAGATGGCGCTGGGCGCAATCGCGCCATTACCGAGTCCGGGCTGGC 868
Db	
556	GGTTCGCCAACGATCAGATGGCGCTGGGCGCAATCGCGCCATTACCGAGTCCGGGCTGGC 497
QY	869 CGTTGGTCCGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATAT 928
Db	
496	CGTTGGTCCGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATAT 437
QY	929 CCGCGGTTAACCCACCATCAACAGGATTTTCGCTGCTGGGGCAACAGCGTGGACCG 988
Db	
436	CCGCGGTTAACCCACCATCAACAGGATTTTCGCTGCTGGGGCAACAGCGTGGACCG 377
QY	989 CTTGCTGCAACTCTCTCAGGCGCAGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACT 1048
Db	
376	CTTGCTGCAACTCTCTCAGGCGCAGCGGTGAAGGGCAATCAGCTGTTGCCCGTCTCACT 317
QY	1049 GGTGAAAAAGAAAAACACCCCTGGCGCCCAATACGCAAAACCGCTCTCCCGCGCGTGGC 1108
Db	
316	GGTGAAGAAAAACACCCCTGGCGCCCAATACGCAAAACCGCTCTCCCGCGCGTGGC 257
QY	1109 CGATTCAATTAATGACGCTGGGACGACAGGTTTCCCGACTGGAAAGCGGCGAGTAGCGCA 1168
Db	
256	CGATTCAATTAATGACGCTGGGACGACAGGTTTCCCGACTGGAAAGCGGCGAGTAGCGCA 197
QY	1169 ACGCAATTAATGAGTTAGCTCACTCAATTAGGCAC 1204
Db	
196	ACGCAATTAATGAGTTAGCTCACTCAATTAGGCAC 161

RESULT 13
 CL655575/c
 LOCUS
 DEFINITION
 CL655575 752 bp DNA linear GSS 09-JUL-2004
 PRI0124a.C01 - PRI0124a.B21 (752) Mixed stage fosmid library of P.
 pacificus var. California Pristionchus pacificus genomic, genomic
 survey sequence.
 CL655575
 G88
 Pristionchus pacificus
 Pristionchus pacificus
 Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
 Neodiplogasteridae; Pristionchus.
 1 (bases 1 to 752)
 Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
 AppADB: an AcedB database for the nematode satellite organism
 Nucleic Acids Res. 32 (1), D421-D422 (2004)
 Contact: Sommer RJ
 Evolutionary Biology
 Max-Planck-Institute for Developmental Biology
 Spemannstr. 37-39, Tuebingen D-72076, Germany
 Tel: 00497071601371
 Fax: 00497071601498
 Email: ralf.sommer@tuebingen.mpg.de
 This library was generated at Caltech, Pasadena, USA and end
 sequenced at Vancouver, Canada.
 Seq primer: 77
 Class: fosmid ends.
 Location/Qualifiers
 1..752
 /organism="Pristionchus pacificus"
 /mol_type="genomic DNA"
 /strain="California"
 /db_xref="taxon:54126"
 /clone_lib="Mixed stage fosmid library of P. pacificus
 var. California"
 /note="Vector: pBpifos-5 Fosmid vector"

ORIGIN

Query Match 18.3%; Score 603; DB 9; Length 752;
 Best Local Similarity 98.9%; Pred. No. 4.7e-167;
 Matches 628; Conservative 0; Mismatches 5; Indels 2; Gaps 2;
 QY 1996 TTCACTGGCCGCTGATTGCTGTGAC -GGGGGTTATGCGTTCAAGTATGAACGGCAA 2054
 DB 752 TTCACTGGCCGCTGATTGCTGTGACGGGGGTTATGCGTTCAAGTATGAACGGCAA 693
 QY 2055 GTACGACATTAAAGCGTGGCGTGGATAACGCTGGCGGAAAGCGGGTCTGACCTTCCT 2114
 DB 692 GTACGACATTAAAGCGTGGCGTGGATAACGCTGGCGGAAAGCGGGTCTGACCTTCCT 633
 QY 2115 GGTGACCTGATTAAACAAACACATGATGCAGACACCGATTACTCCATCGCAGAAC 2174
 DB 632 GGTGACCTGATTAAACAAACACATGATGCAGACACCGATTACTCCATCGCAGAAC 573
 QY 2175 TGCCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCGCTGGGCATGGTCCCAACAT 2234
 DB 572 TGCCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCGCTGGGCATGGTCCCAACAT 513
 QY 2235 CGACACCGAAGTGAATTATGGTGTAAACGGTACTGCCGACCTTCAAGGGTCAACCATC 2294
 DB 512 CGACACCGAAGTGAATTATGGTGTAAACGGTACTGCCGACCTTCAAGGGTCAACCATC 453
 QY 2295 CAACCCGTTGTTGGCGTGTGAGCGAGGTATTAAACGGCCAGTCCGCAACAGAGCT 2354
 DB 452 CAACCCGTTGTTGGCGTGTGAGCGAGGTATTAAACGGCCAGTCCGCAACAGAGCT 393
 QY 2355 GCGAAAGAGTTCCTCGAAACCTATCTGTGATGAAGGTCTGGAAGCGGTTAA-TA 2413
 DB 392 GCGAAAGAGTTCCTCGAAACCTATCTGTGATGAAGGTCTGGAAGCGGTTAATTA 333

QY 2414 AAGACAAACCGCTGGTGGCTAGCGTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATC 2473
 DB 332 AGGCCAACCCGCTTGGTGGCTAGCGTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATC 273
 QY 2474 CACGATTTCGGCCACCATGAAACCCGAGAAAGTGAAATCATCCGAAACATCCCGC 2533
 DB 272 CACGATTTCGGCCACCATGAAACCCGAGAAAGTGAAATCATCCGAAACATCCCGC 213
 QY 2534 AGATGTCGGCTTTCTGCTATGCGGTGCTACTGCGGTGATCAACGCCGCCGCGGTGTC 2593
 DB 212 AGATGTCGGCTTTCTGCTATGCGGTGCTACTGCGGTGATCAACGCCGCCGCGGTGTC 153
 QY 2594 AGATGTCGATGAAGCCCTGAAAGACGCGGACT 2628
 DB 152 AGATGTCGATGAAGCCCTGAAAGACGCGGACT 118
 RESULT 14
 AY080094/c
 LOCUS
 DEFINITION
 AY080094 608 bp DNA linear GSS 06-NOV-2002
 uncultured marine virus genomic clone SIO51p3D4L, genomic survey
 sequence.
 AY080094
 AY080094.1 GI:24745276
 GSS.
 uncultured marine virus
 uncultured marine virus
 Viruses; environmental samples.
 1 (bases 1 to 608)
 Breitbart,M., Salamon,P., Andresen,B., Mahaffy,J.M., Segall,A.M.,
 Mead,D., Azam,F. and Rohwer,F.
 Genomic analysis of uncultured marine viral communities
 Proc. Natl. Acad. Sci. U.S.A. 99 (22), 14250-14255 (2002)
 22294988
 12384570
 Contact: Rohwer F
 Biology Dept.
 San Diego State University
 5500 Campanile Dr, San Diego, CA 92102, USA
 Tel: 6195941336
 Fax: 619595676
 Email: forest@sunstroke.sdsu.edu
 Class: shotgun.
 Location/Qualifiers
 1..608
 /organism="uncultured marine virus"
 /mol_type="genomic DNA"
 /db_xref="taxon:186617"
 /clone="SIO51p3D4L"
 /clone_lib="Scripps Pier (La Jolla, CA) uncultured virus
 community"
 /note="Marine viruses were isolated from 200 liters of
 surface seawater using a combination of differential
 filtration and density-dependent gradient centrifugation.
 Linker-amplified shotgun libraries were constructed by
 randomly shearing the total marine viral community DNA,
 end-repairing, ligating dsDNA linkers to the ends, and
 amplifying the fragments using Vent DNA polymerase. The
 resulting fragments were ligated into the pSMART vector
 and electroporated into MC12 cells (Lucigen; Middleton,
 WI)."

ORIGIN

Query Match 17.9%; Score 589.8; DB 8; Length 608;
 Best Local Similarity 99.7%; Pred. No. 3.7e-163;
 Matches 591; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 287 CAGTCTGGCCCTGCACGCGCCCTGCCAAATTCGCGGCGATTAAATCTCGCGCGATCA 346
 DB 608 CAGTCTGGCCCTGCACGCGCCCTGCCAAATTCGCGGCGATTAAATCTCGCGCGATCA 549

```
QY 347 ACTGGGTGCCACGGTGGTGGTGCATGGTAGAAGCAAGCGCGTCGAAGCCTGTAAAGC 406
Db 548 ACTGGGTGCCACGGTGGTGGTGCATGGTAGAAGCAAGCGCGTCGAAGCCTGTAAAGC 489
QY 407 GCGGTGCACAAATCTTCGCGCAACAGCGTCAGTGGGCTGATCAATTAATATCCGCTGGA 466
Db 488 GCGGTGCACAAATCTTCGCGCAACAGCGTCAGTGGGCTGATCAATTAATATCCGCTGGA 429
QY 467 TGACCAAGATGCCATGCTGTGGAGAGCTGCCTGCACTAATGTTCGGGCTTATTTCTTGA 526
Db 428 TGACCAAGATGCCATGCTGTGGAGAGCTGCCTGCACTAATGTTCGAGCGTTATTTCTTGA 369
QY 527 TGTCTCTGACCAAGACACCAATCAACAGATATATTTCTCCCATGAAGCGGTACGGGACT 586
Db 368 TGTCTCTGACCAAGACACCAATCAACAGATATATTTCTCCCATGAAGCGGTACGGGACT 309
QY 587 GGGCGTGAAGCATCTGGTTCGATTCGGTCCACAGCAAAATCGCGCTGTAGCGGCGCCATT 646
Db 308 GGGCGTGAAGCATCTGGTTCGATTCGGTCCACAGCAAAATCGCGCTGTAGCGGCGCCATT 249
QY 647 AAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGGTCATAAATATCTACTCGCAATCA 706
Db 248 AAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGGTCATAAATATCTACTCGCAATCA 189
QY 707 AATTACGCGGATAGCGGAACGGAAGGAGCTGGAGTGCATGTCGGGTTTTCACAAAC 766
Db 188 AATTACGCGGATAGCGGAACGGAAGGAGCTGGAGTGCATGTCGGGTTTTCACAAAC 129
QY 767 CATGCAATGCTGAATGAGGCGCATCGTTCCCACTCGCATGCTGGTTGCCAAGATCAGAT 826
Db 128 CATGCAATGCTGAATGAGGCGCATCGTTCCCACTCGCATGCTGGTTGCCAAGATCAGAT 69
QY 827 GCGCTGGGCGCAATGCGCGCATTAACCGAGTCCGGGCTCGCGTGTGGTGGCG 879
Db 68 GCGCTGGGCGCAATGCGCGCATTAACCGAGTCCGGGCTCGCGTGTGGTGGCG 16
```

```
RESULT 15
AL037742/c
LOCUS DKF2p564C157.r1.564 (synonym: hfbz2) Homo sapiens cdna clone
DEFINITION DKF2p564C157.5, mRNA sequence.
ACCESSION AL037742.1 GI:5407088
VERSION 1 (bases 1 to 589)
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 589)
AUTHORS Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
TITLE EST (Bloecker, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
```

```
FEATURES
source
1..589
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKF2p564C157"
/organism="Homo sapiens"

```

This clone (DKF2p564C157) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers

```
/tissue_type="brain"
/dev_stage="fetal"
/lab_host="X1-2blue"
/clone_lib="564 (synonym: hfbz2)"
/note="Vector: pAMP1; Site_1: NotI; Site_2: SalI"
```

ORIGIN

```
Query Match 17.8%; Score 586.4; DB 1; Length 589;
Best Local Similarity 99.7%; Pred. No. 3.7e-162;
Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 517 TATTTCTTTGATGCTCTCGACAGACACCCATCAACAGTATATTTTCTCCCATGAAGAGC 576
Db 589 TATTTCTTTGATGCTCTCGACAGACACCCATCAACAGTATATTTTCTCCCATGAAGAGC 530
QY 577 GTAACGCACTGGGCGTGAGCATCTGTGCGATTTGGGTCACAGCAAAATCGCGCTGTTAG 636
Db 529 GTACGCGACTGGGCGTGAGCATCTGTGCGATTTGGGTCACAGCAAAATCGCGCTGTTAG 470
QY 637 CGGGCCCAATTAAGTTCTCTCGGGCGCTCTGCGTCTGGCTGGGTCATAAATATCTCA 696
Db 469 CGGGCCCAATTAAGTTCTCTCGGGCGCTCTGCGTCTGGCTGGGTCATAAATATCTCA 410
QY 697 CTCGCAATCAAAATTCAGCCGATAGCGBAACGGGAAGCGACTGGAGTGCCATGTCGGGT 756
Db 409 CTCGCAATCAAAATTCAGCCGATAGCGBAACGGGAAGCGACTGGAGTGCCATGTCGGGT 350
QY 757 TTCAACAAACCATGCAAAATGCTGAATGAGGCGCATCGTTCCCACTCGCATGCTGGTTGCCA 816
Db 349 TTCAACAAACCATGCAAAATGCTGAATGAGGCGATCGTTCCCACTCGCATGCTGGTTGCCA 290
QY 817 ACGATCAGATGGGCTGGGCGCAATTCGCGCATTAACCGAGTCCGGGCTCGCGTGTGGTG 876
Db 289 ACGATCAGATGGGCTGGGCGCAATTCGCGCATTAACCGAGTCCGGGCTCGCGTGTGGTG 230
QY 877 CGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCGT 936
Db 229 CGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCGT 170
QY 937 TAACCAACCATCAAAACAGGATTTTTCGCTGCTGGGCGCAAAACAGCGTGGACCGCTGCTGC 996
Db 169 TAACCAACCATCAAAACAGGATTTTTCGCTGCTGGGCGCAAAACAGCGTGGACCGCTGCTGC 110
QY 997 AACTCTCTCAGGGCCAGCGCGTGAAGGGCAATCAGCTGTTGCCCGCTCTCACCTGGTGA 1056
Db 109 AACTCTCTCAGGGCCAGCGCGTGAAGGGCAATCAGCTGTTGCCCGCTCTCACCTGGTGA 50
QY 1057 GAAAAACCAACCTCGCGCCCAATACGCAACCGCTCTCTCCCGCGCGTT 1105
Db 49 GAAAAACCAACCTCGCGCCCAATACGCAACCGCTCTCTCCCGCGCGTT 1
```

Search completed: August 3, 2005, 02:02:46
Job time : 6748.2 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 1309.2 Seconds
(without alignments)
16314.373 Million cell updates/sec

Title: US-09-765-555B-14
Perfect score: 3300
Sequence: 1 ccgacacatcgatggtgc.....acgacgttcggactaagcgt 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq:*
- 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3300	100.0	3300	10	US-09-765-555-14
2	3300	100.0	3300	10	US-09-765-555-15
3	3226.4	97.8	3300	10	US-09-765-555-18
4	3215.2	97.4	3300	10	US-09-765-555-17
5	3207.2	97.2	3300	10	US-09-765-555-16
6	2705.8	82.0	7373	18	US-10-257-384A-1
7	2705.8	82.0	8101	17	US-10-149-472-5
8	2701	81.8	2701	19	US-10-343-859-8
9	2698	81.8	2698	19	US-10-343-859-9
10	2690.4	81.5	2690.4	18	US-10-263-153-40
11	2690.4	81.5	2690.4	18	US-10-263-153-30
12	2690.4	81.5	2690.4	18	US-10-263-153-35
13	2690.4	81.5	2690.4	18	US-10-263-153-25
14	2690.4	81.5	2690.4	18	US-10-263-153-20
15	2690.4	81.5	2690.4	18	US-10-263-153-61
16	2690.4	81.5	2690.4	18	US-10-263-153-66
17	2690.4	81.5	2690.4	18	US-10-263-153-71
18	2690.4	81.5	2690.4	18	US-10-263-153-15
19	2690.4	81.5	2690.4	18	US-10-263-153-10
20	2690.4	81.5	2690.4	18	US-10-263-153-3
21	2636.8	79.9	2636.8	10	US-09-908-943A-194
22	2636.8	79.9	2636.8	10	US-10-801-487-194
23	2636.8	79.9	2636.8	20	US-10-801-938-194
24	2636.8	79.9	2636.8	20	US-10-801-509-194
25	2636.8	79.9	2636.8	20	US-10-801-486-194
26	2636.8	79.9	2636.8	21	US-10-801-493-194
27	2605.4	79.0	2605.4	18	US-10-263-153-7
28	1609.4	48.8	1609.4	16	US-10-241-596-137
29	1474	44.7	1474	17	US-10-313-963A-54
30	1462.4	44.3	1462.4	21	US-10-495-491-2
31	1462.4	44.3	1462.4	21	US-10-495-491-1
32	1462.4	44.3	1462.4	21	US-10-495-491-3
33	1462.4	44.3	1462.4	21	US-10-495-491-4
34	1417.6	43.0	1417.6	10	US-09-987-763-30
35	1417.6	43.0	1417.6	13	US-10-024-809-3
36	1238.4	37.5	1238.4	17	US-10-313-963A-52
37	1238.4	37.5	1238.4	21	US-10-671-995A-2
38	1238.4	37.5	1238.4	21	US-10-671-995A-1
39	1237.2	37.5	1237.2	10	US-09-995-898A-32
40	1237.2	37.5	1237.2	17	US-10-420-034A-32
41	1237.2	37.5	1237.2	21	US-10-855-897-2
42	1204	36.5	1204	17	US-10-457-372-5
43	1204	36.5	1204	17	US-10-270-176-1
44	1204	36.5	1204	18	US-10-270-176-2
45	1204	36.5	1204	18	US-10-270-176-19

ALIGNMENTS

RESULT 1
US-09-765-555-14
; Sequence 14, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m1 and zinc finger
; OTHER INFORMATION: Protein ZFPm1
US-09-765-555-14

Query Match 100.0%; Score 3300; DB 10; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCACCATCGAATGGTGC AAAACCTTTCCGGGTATGCGCATGATAGCGCCGGGAAGAGA 60
Db CGCACCATCGAATGGTGC AAAACCTTTCCGGGTATGCGCATGATAGCGCCGGGAAGAGA 60
QY 61 GTC AATTCAGGGTGTGAATGTGA AACCAAGTAA CTTTATACGATGTCCGAGATGCGG 120
Db GTC AATTCAGGGTGTGAATGTGA AACCAAGTAA CTTTATACGATGTCCGAGATGCGG 120
QY 121 GTGTCCTCTTATCAGACCGTTTCCCGCGTGGTGAA CCGAGCGCAGCCACGCTTTCTCGAAAA 180
Db GTGTCCTCTTATCAGACCGTTTCCCGCGTGGTGAA CCGAGCGCAGCCACGCTTTCTCGAAAA 180
QY 181 CGCGGAAAAAGTGAAGCGCGCATGGCGAGCTGAATTACATTTCCCAACCGCGTGGCAC 240
Db CGCGGAAAAAGTGAAGCGCGCATGGCGAGCTGAATTACATTTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAACAGTCGTTGCTGATTTGGGTTGGCACTTCCAGTCTGGCCCTGC 300
Db AACAACTGGCGGGCAACAGTCGTTGCTGATTTGGGTTGGCACTTCCAGTCTGGCCCTGC 300
QY 301 ACGGCGCGTCGCAAAATTTGTCGGCGGATTAATAATCTCGGCGCGATCAACTGGGTGCCAGCG 360
Db ACGGCGCGTCGCAAAATTTGTCGGCGGATTAATAATCTCGGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTGTGATGGTAGAACGAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCACAATC 420
Db TGGTGGTGTGATGGTAGAACGAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACGGTCAGTGGGCTGATCATTAATCTATCCGCTGGATGACCAAGATGCCA 480
Db TTCTCGCGCAACGGTCAGTGGGCTGATCATTAATCTATCCGCTGGATGACCAAGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTCAGCCAGA 540
Db TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTCAGCCAGA 540
QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Db CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
QY 601 TGGTGCATTTGGGTACCAAGCAATCGGCTGTTAGCGGGCCCATTAAGTTCCTGCTCGG 660
Db TGGTGCATTTGGGTACCAAGCAATCGGCTGTTAGCGGGCCCATTAAGTTCCTGCTCGG 660
QY 661 CGCGTCTCGGCTCGGCTGGCATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db CGCGTCTCGGCTCGGCTGGCATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGCGCATGGAGTGCCATGTCGGGTTTTCACAAACCATGCAATGCTGA 780
Db CGGAACGGGAAGCGCATGGAGTGCCATGTCGGGTTTTCACAAACCATGCAATGCTGA 780
QY 781 ATGAGGCGATCGTTCCCACTCGCATGCTGGTTCGAACGATCAGATGGCGTGGCGCAA 840
Db ATGAGGCGATCGTTCCCACTCGCATGCTGGTTCGAACGATCAGATGGCGTGGCGCAA 840
QY 841 TGC CGCATTACCGAGTCCGGGCTGGCGGTGGTGCGGATATCTCGTAGTGGGATACG 900
Db TGC CGCATTACCGAGTCCGGGCTGGCGGTGGTGCGGATATCTCGTAGTGGGATACG 900
QY 901 ACGATACGAAGAAGCTCATGTTATATCCGCGCGTTAA CCAATCAAAACAGGATTTTC 960
Db ACGATACGAAGAAGCTCATGTTATATCCGCGCGTTAA CCAATCAAAACAGGATTTTC 960
QY 961 GCCTGTGGGCGAAACAGCGTGCACGCTTGCTGCAACTCTCAGGCGCCAGCGGTGA 1020
Db GCCTGTGGGCGAAACAGCGTGCACGCTTGCTGCAACTCTCAGGCGCCAGCGGTGA 1020
QY 1021 AGGCAATACGCTGTTCCCGTCTCACTGTGTGAAGAAAAACCA CCGCTGGCGCCCAATA 1080
Db AGGCAATACGCTGTTCCCGTCTCACTGTGTGAAGAAAAACCA CCGCTGGCGCCCAATA 1080
QY 1081 CGCAACCGCGCTCTCCCGCGCGTTGGCGGATTCATTAATGCACTGGCAGCAGAGTTT 1140

Db CGCAACCGCGCTCTCCCGCGCGTTGGCGGATTCATTAATGCACTGGCAGCAGAGTTT 1140
QY 1141 CCCGACTGGAAAGCGGCGAFTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
Db CCCGACTGGAAAGCGGCGAFTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTTGACAGCTTATCATGCACTGACCGTGACCGTGCACCAATGCTTCTGGCG 1260
Db GCACAAATCTCATGTTTTGACAGCTTATCATGCACTGACCGTGACCGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGCAAGCCATCGAAAGCTGTGATGCGTGTGACAGGTGTAATAATCACTGCATAAATTCG 1320
Db TCAGGCAAGCCATCGAAAGCTGTGATGCGTGTGACAGGTGTAATAATCACTGCATAAATTCG 1320
QY 1321 TGTGCTCAAGGGCGACTCCCGTTCTGGATAATGTTTTTTTGGCCGACATCATAAACGGTT 1380
Db TGTGCTCAAGGGCGACTCCCGTTCTGGATAATGTTTTTTTGGCCGACATCATAAACGGTT 1380
QY 1381 CTGGCAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTTGGGA 1440
Db CTGGCAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTTGGGA 1440
QY 1441 ATTGTGAGCGGATAACAAATTTTCACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCACGA 1500
Db ATTGTGAGCGGATAACAAATTTTCACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCACGA 1500
QY 1501 GCACCTCACCAACAGCAACCATAGATTATGAAAACCTGAAAGGATAAATCTGGTAATCTGG 1560
Db GCACCTCACCAACAGCAACCATAGATTATGAAAACCTGAAAGGATAAATCTGGTAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGCTCTCGCTGCAAGTCGGTAAGAAATTCGGAAGAT 1620
Db ATTAACGGCGATAAAGGCTATAACGCTCTCGCTGCAAGTCGGTAAGAAATTCGGAAGAT 1620
QY 1621 ACCGGAATTAAGTACCGTTGAGCATCCCGATTAACCTGGAAGAGAAATTTCCCAACAGGTT 1680
Db ACCGGAATTAAGTACCGTTGAGCATCCCGATTAACCTGGAAGAGAAATTTCCCAACAGGTT 1680
QY 1681 GCGGCAACTGGCATGGCCCTGACATTAATCTTCTGGGCA CACGACCGCTTTGGTGGCTAC 1740
Db GCGGCAACTGGCATGGCCCTGACATTAATCTTCTGGGCA CACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTCGCCTGTTGGCTGAAATCACCCCGA CAAAGCGTTCCAGGACAAGCTGTAT 1800
Db GCTCAATCTCGCCTGTTGGCTGAAATCACCCCGA CAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGTACGTTACACGCGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
Db CCGTTTACCTGGGATGCGTACGTTACACGCGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATACAAAGATCTCGTGGCGAACCCGCCCAAAAACCTGGGAA 1920
Db GAAGCGTTATCGCTGATTATACAAAGATCTCGTGGCGAACCCGCCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCAAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
Db GAGATCCCGCGCTGGATAAAGAACTGAAAGCAAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGGTACTTCACTGGCGCTGATTGCTGTGACGCGGGGTTATGCGTTCAAG 2040
Db CTGCAAGAACCGGTACTTCACTGGCGCTGATTGCTGTGACGCGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACGGCAAGTACGACATTAAGAGCTGGCGTGGATAAACGCTGGCGGAAGCG 2100
Db TATGAAACGGCAAGTACGACATTAAGAGCTGGCGTGGATAAACGCTGGCGGAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTTAAAAACAAACATGAATGCAGACACCGATTAC 2160
Db GGTCTGACCTTCTCGTTGACCTGATTTAAAAACAAACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAACGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCGCCGTGG 2220

```
Db 2161 TCCATCGCAGAACTGCGCTTTTAATAAAGCGGAAACAGCGATGACCATCAACGGCCCGGTGG 2220
Qy 2221 GCATGGTCCAAATCGACACACGAAAGTGAATTAATGGTGAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACACGAAAGTGAATTAATGGTGAACGGTACTGCGGACCTTC 2280
Qy 2281 AAGGTCACCATCAACACCGTTCTGTGGCGTGTGAGGGCAGGTATTAAACGCCGCCAGT 2340
Db 2281 AAGGTCACCATCAACACCGTTCTGTGGCGTGTGAGGGCAGGTATTAAACGCCGCCAGT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAAAGAGTCTTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTCTTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGGTAAATAAGACAAACCGTGGTGGTCCGTAGCGCTGAAGTCTTACAGAAAG 2460
Db 2401 GAAGCGGTAAATAAGACAAACCGTGGTGGTCCGTAGCGCTGAAGTCTTACAGAAAG 2460
Qy 2461 TTGGCGAAAGATCCACGTAATTCGCGCACCATGGAAGAGCCGAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTAATTCGCGCACCATGGAAGAGCCGAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCGACAGATGTCGCTTCTGGTATGCGTGTGCTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACATCCGACAGATGTCGCTTCTGGTATGCGTGTGCTGCTGCTGATCAACGCC 2580
Qy 2581 GCCAGCGGTGCTGACAGTGTGATGAAGCCCTGAAAGAGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGCTGACAGTGTGATGAAGCCCTGAAAGAGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACACACAAACAAATAACAAACCTCTGGGATGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACACACAAACAAATAACAAACCTCTGGGATGAGGGAAGGATTCAGAAATTC 2700
Qy 2701 GGATCCTCTCTCTGTGGCCAGCGGCCCTCGAGCCGCGGAGAGCCCTATGCTGT 2760
Db 2701 GGATCCTCTCTCTGTGGCCAGCGGCCCTCGAGCCGCGGAGAGCCCTATGCTGT 2760
Qy 2761 CCGGAATGTGGTAAGTCTTCTCTCAGAGCTCTCACCTGCTGCGCACCCAGCGTACCCAC 2820
Db 2761 CCGGAATGTGGTAAGTCTTCTCTCAGAGCTCTCACCTGCTGCGCACCCAGCGTACCCAC 2820
Qy 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGGCAATCTTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGGCAATCTTTTAGCCAGTCCAGCAAC 2880
Qy 2881 CTGGTGGCCATCAAGCACTCATCTGCGAGAGAGCCATACAAATGTCCAGATGTGGC 2940
Db 2881 CTGGTGGCCATCAAGCACTCATCTGCGAGAGAGCCATACAAATGTCCAGATGTGGC 2940
Qy 2941 AAGTCTTTCTCTCGGTCTGACAACTCTGTCGGCACCAACGTACTCACACCGGGGAGAG 3000
Db 2941 AAGTCTTTCTCTCGGTCTGACAACTCTGTCGGCACCAACGTACTCACACCGGGGAGAG 3000
Qy 3001 CCTATGCTTGTCCGGAATGTGGTAAGTCTTTCAGCCGAGCGATAAATCTGTGGCCGAC 3060
Db 3001 CCTATGCTTGTCCGGAATGTGGTAAGTCTTTCAGCCGAGCGATAAATCTGTGGCCGAC 3060
Qy 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAAATGCCAGAGTGGGCAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAAATGCCAGAGTGGGCAATCTTTTAGC 3120
Qy 3121 CAGGCGGCGCACTGCGCAGCCATCAACGCACTCTACTCTGGGAGAGGCAATCAAAATGT 3180
Db 3121 CAGGCGGCGCACTGCGCAGCCATCAACGCACTCTACTCTGGGAGAGGCAATCAAAATGT 3180
Qy 3181 CCAGAAATGTGGCAAGTCTTCTCTCGGTCTGCAATCTCTCGTCCGCGACCAACGTAATC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTCTCTCGGTCTGCAATCTCTCGTCCGCGACCAACGTAATC 3240
Qy 3241 ACCGGTAAAAAATAGTGGCCAGGCGCGCGAGTACCGGTAGCAAGTTCGGGACTACGCT 3300
Db 3241 ACCGGTAAAAAATAGTGGCCAGGCGCGCGAGTACCGGTAGCAAGTTCGGGACTACGCT 3300
```

RESULT 2

```
US-09-765-555-15
; Sequence 15, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE REFERENCE: 27801-20014, 40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m2 and zinc finger
; OTHER INFORMATION: protein ZFPm2
US-09-765-555-15
```

```
Query Match 100.0%; Score 3300; DB 10; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGCAGAAACCTTTTCGCGGTATGGCATGATAGGCCCGGAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAGAAACCTTTTCGCGGTATGGCATGATAGGCCCGGAGAGA 60
Qy 61 GTCATTCAGGGTGTGAATGTGAACCAAGTAACTTATACGATGTCGAGAGTATGCG 120
Db 61 GTCATTCAGGGTGTGAATGTGAACCAAGTAACTTATACGATGTCGAGAGTATGCG 120
Qy 121 GTGTCCTTATCAGACCGTTTCCGCGGTGTAACCAAGCCAGCCACGCTTTCTGCGAAA 180
Db 121 GTGTCCTTATCAGACCGTTTCCGCGGTGTAACCAAGCCAGCCACGCTTTCTGCGAAA 180
Qy 181 CCGGGAAGAAAGTGAAGCGCGATGCGGAGCTGAATTAATTCACCAACCGCTGGCAC 240
Db 181 CCGGGAAGAAAGTGAAGCGCGATGCGGAGCTGAATTAATTCACCAACCGCTGGCAC 240
Qy 241 AACAACTGGCGGCAACAGTCTGTTGCTGATGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAACAGTCTGTTGCTGATGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGGCGCTGCGCAATTCGCGCGGATTAATCTCGCGCGATCAATCTGGGTGCGCAGG 360
Db 301 ACGGCGCTGCGCAATTCGCGCGGATTAATCTCGCGCGATCAATCTGGGTGCGCAGG 360
Qy 361 TGGTGTGTCGATGTAGAACGAAGCGCGTGAAGCCCTGTAAAGCGCGGTGCAAAATC 420
Db 361 TGGTGTGTCGATGTAGAACGAAGCGCGTGAAGCCCTGTAAAGCGCGGTGCAAAATC 420
Qy 421 TTCTCGCGCAACGCGTCAAGTGGGTGATCAATTAATCTATCGCTGATGACAGGATGCCA 480
Db 421 TTCTCGCGCAACGCGTCAAGTGGGTGATCAATTAATCTATCGCTGATGACAGGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCTGCAATGTTCCGGCGTATTTCTGATGCTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCAATGTTCCGGCGTATTTCTGATGCTCTGACCCAGA 540
Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTGCATTTGGGTCAACGCAAAATCGCGCTGTTAGCGGGGCCCATTTAAAGTCTGTCTCGG 660
```


Db 601 |||||TGCTCGCATTTGGTTCACCAAGCAATTCGGCTGTAGCGGGCCCAITTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTGGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db 661 CGCGTCTGGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
QY 721 CGGAA CGGAAAGCGACATGGAGTGCATGTCCGGTTCCTCAACAAACCAATGCAAAATGCTGA 780
Db 721 CGGAAAGCGGACATGGAGTGCATGTCCGGTTCCTCAACAAACCAATGCAAAATGCTGA 780
QY 781 ATGAGGGCATGTTCCGACCTGGATGCTGGTTGGCCAAAGATCAGATGGCGTGGCGCAA 840
Db 781 ATGAGGGCATGTTCCGACCTGGATGCTGGTTGGCCAAAGATCAGATGGCGTGGCGCAA 840
QY 841 TCGCGGCCATTAACCGAGTCCGGCTGGCGTGGTGGTGGGATATCTCGTATGGGATAG 900
Db 841 TCGCGGCCATTAACCGAGTCCGGCTGGCGTGGTGGGATATCTCGTATGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCTTAACCAACCAATCAAAAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCTTAACCAACCAATCAAAAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACGAGCTGGAACCGTTCGTCGAATCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGCTGGAACCGTTCGTCGAATCTCTCAGGGCCAGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGGCGCTCTCACTGGTGAAGAAACCAACCTTCGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCGCTCTCACTGGTGAAGAAACCAACCTTCGCGCCCAATA 1080
QY 1081 CGCAACCGCTCTCCCGCGCTTGGCGATTCATTAATGCAGCTGSCACGACAGGTTT 1140
Db 1081 CGCAACCGCTCTCCCGCGCTTGGCGATTCATTAATGCAGCTGSCACGACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGAATTAATGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGAATTAATGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACCGTGCACCAATCTTCTGCG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACCGTGCACCAATCTTCTGCG 1260
QY 1261 TCAGGACCGCATCGAAGCTGTGATGCTGTGAGGTGTAATCACTGCAATAATTCG 1320
Db 1261 TCAGGACCGCATCGAAGCTGTGATGCTGTGAGGTGTAATCACTGCAATAATTCG 1320
QY 1321 TGTGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCAACCGTT 1380
Db 1321 TGTGCTCAAGCGCACCTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCAACCGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
QY 1441 ATTGTGAGCGGATAACAAATTTACAGAGAAACAGCAGTCCGTTTGGTGTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTTACAGAGAAACAGCAGTCCGTTTGGTGTTCACGA 1500
QY 1501 GCACCTTCACCAACAGGACCATGATTAATGAACCTGAAGGTAACCTGGTAATCTGG 1560
Db 1501 GCACCTTCACCAACAGGACCATGATTAATGAACCTGAAGGTAACCTGGTAATCTGG 1560
QY 1561 ATTAACCGCGATAAAGCTATAACCGTCTCGCTGAAGTCCGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACCGCGATAAAGCTATAACCGTCTCGCTGAAGTCCGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAAGTTGACCTCCGATTAACCTGGAAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAAGTTGACCTCCGATTAACCTGGAAGAAATTCACAGGTT 1680
QY 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740

Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCAACCCCGACAAAGGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCAACCCCGACAAAGGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACGGCAAGCTGATGCTTACCGGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACGGCAAGCTGATGCTTACCGGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATTAACAAAGATCTGTCGCCGAACCCGCCAAAACCTCGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATTAACAAAGATCTGTCGCCGAACCCGCCAAAACCTCGGAA 1920
QY 1921 GAGATCCGGCGCTGGATTAAGAACTGAAAGCGGAAAGGTAAGAGCGCGCTGATTTCAAC 1980
Db 1921 GAGATCCGGCGCTGGATTAAGAACTGAAAGCGGAAAGGTAAGAGCGCGCTGATTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTTGCTGCTGACGGGGTTATGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTTGCTGCTGACGGGGTTATGGTTCAAG 2040
QY 2041 TATGAAACCGGCAAGTACGACATTTAAAGACGTTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGGCAAGTACGACATTTAAAGACGTTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGTTGACCTGATTTAAACAAACACATGAATGCGACACCGATTTAC 2160
Db 2101 GGTCTGACCTTCTGTTGACCTGATTTAAACAAACACATGAATGCGACACCGATTTAC 2160
QY 2161 TCCATCCGAGAAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCCCGTGG 2220
Db 2161 TCCATCCGAGAAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCCCGTGG 2220
QY 2221 GCATGTGCTCAACATCGACACAGCAAGTGAATTTATGGTGAACGGTACTGCGACCTTC 2280
Db 2221 GCATGTGCTCAACATCGACACAGCAAGTGAATTTATGGTGAACGGTACTGCGACCTTC 2280
QY 2281 AAGGGTCAACATTCGCAACCGTTTGGGTGCTGAGCGCAGGTATTAACGCCCGCAGT 2340
Db 2281 AAGGGTCAACATTCGCAACCGTTTGGGTGCTGAGCGCAGGTATTAACGCCCGCAGT 2340
QY 2341 CCAGAACAAAGAGCTGGCAAAAGAGTTTCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCAGAACAAAGAGCTGGCAAAAGAGTTTCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
QY 2401 GAAGCGGTTAATAAAGACAAAACCGTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAAACCGTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCAGTATTTGCCGCCACCATGGAACCGCCAGAAAGGTGAATCAGT 2520
Db 2461 TTGGCGAAAGATCCAGTATTTGCCGCCACCATGGAACCGCCAGAAAGGTGAATCAGT 2520
QY 2521 CCAGAACATCCCGCAGATGTCCGCTTTCTGGTATGCGGTGCTGCTGCTGATGCAACGCC 2580
Db 2521 CCAGAACATCCCGCAGATGTCCGCTTTCTGGTATGCGGTGCTGCTGCTGATGCAACGCC 2580
QY 2581 GCAGCGGTGCTGAGACTGTGCAATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCAGCGGTGCTGAGACTGTGCAATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACACAAACAAATTAACAAATTAACAACTCGGGATCGAGGAGGATTTTCAGAAATTC 2700
Db 2641 AACACAAACAAATTAACAAATTAACAACTCGGGATCGAGGAGGATTTTCAGAAATTC 2700
QY 2701 GGAATCTTCTCTGTGTCGCGCGCTTCGAGCCCGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGAATCTTCTCTGTGTCGCGCGCTTCGAGCCCGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTTGTAAGTCTTCTCTCAGAGCTCTCACTGGTGGCCACAGCGTATCCAC 2820
Db 2761 CCGGAATGTTGTAAGTCTTCTCTCAGAGCTCTCACTGGTGGCCACAGCGTATCCAC 2820

```
QY 2821 ACCGGTGAACCAACCGTATATAATGCCAGAGTCGGCGCAAAATCTTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACCGGTGAACCAACCGTATATAATGCCAGAGTCGGCGCAAAATCTTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGGTGGCCCATCAACGACCTCATACTGCGGAGAGCCATACAAATGTCCAGATGTGGC 2940
Db 2881 CTGGTGGCCCATCAACGACCTCATACTGCGGAGAGCCATACAAATGTCCAGATGTGGC 2940
QY 2941 AAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTAATCTCACACCGGGGAGAG 3000
Db 2941 AAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTAATCTCACACCGGGGAGAG 3000
QY 3001 CCCTATGCTTGTCCGGAATGTGGTAAGTCTTTTCAAGTCCTTTCAAGCGCAGGCAATAAATCTGGTGGCCAC 3060
Db 3001 CCCTATGCTTGTCCGGAATGTGGTAAGTCTTTTCAAGTCCTTTCAAGCGCAGGCAATAAATCTGGTGGCCAC 3060
QY 3061 CAGCGTACCCACACCGGTGAAAAACCGTATATAATGCCAGAGTCGGCGCAAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACCGGTGAAAAACCGTATATAATGCCAGAGTCGGCGCAAAATCTTTTAGC 3120
QY 3121 CAGCGCGGCCACTCGCCAGCCATCAAGCCACTCATCTGCGGAGAGCCATACAAATGT 3180
Db 3121 CAGCGCGGCCACTCGCCAGCCATCAAGCCACTCATCTGCGGAGAGCCATACAAATGT 3180
QY 3181 CCAGAAATGTGGCAAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTAATCTCAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTTCTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTAATCTCAC 3240
QY 3241 ACCGGTAAAAAATAGTGGCCAGCGCCGACAGTACCGGTACAGCTTCCGAGCTACCGCT 3300
Db 3241 ACCGGTAAAAAATAGTGGCCAGCGCCGACAGTACCGGTACAGCTTCCGAGCTACCGCT 3300
```

RESULT 3

```
US-09-765-555-18
; Sequence 18, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-Ap3 and zinc finger
; OTHER INFORMATION: protein ZFPap3
US-09-765-555-18
```

```
Query Match 97.8%; Score 3226.4; DB 10; Length 3300;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3254; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 1 CCGACACCATCAATGGTGCAGAACTTTTCGGGTATGCGATGATAGCCCGGAGAGA 60
Db 1 CCGACACCATCAATGGTGCAGAACTTTTCGGGTATGCGATGATAGCCCGGAGAGA 60
QY 61 GTCAATTACAGGTGGTGAATGTGAACACCAAGTAACTATACGATGTGCGAGAGTATGCCG 120
Db 61 GTCAATTACAGGTGGTGAATGTGAACACCAAGTAACTATACGATGTGCGAGAGTATGCCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCACGCTTTCTGCGAAAA 180
```

```
Db 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCACGCTTTCTGCGAAAA 180
QY 181 CGCGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTAATATCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTAATATCCCAACCGCGTGGCAC 240
QY 241 AACAACTGCGCGGCAAAACAGTCTGTTGCTGATTTGGCGTTCACACTCTGAGTCTGGCCCTGC 300
Db 241 AACAACTGCGCGGCAAAACAGTCTGTTGCTGATTTGGCGTTCACACTCTGAGTCTGGCCCTGC 300
QY 301 ACGGCGCGTCCCAAAATTCGCGCGGATTAATCTCGCGCGATCAAACTGGGTGCGCAGCG 360
Db 301 ACGGCGCGTCCCAAAATTCGCGCGGATTAATCTCGCGCGATCAAACTGGGTGCGCAGCG 360
QY 361 TGGTGGTTCGATGATAGCAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCAACAATC 420
Db 361 TGGTGGTTCGATGATAGCAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCAACAATC 420
QY 421 TTCTCGCGCAACGGGTGATGCGGTGATCAATTAATATCCGCTGATGACACAGGATGCCA 480
Db 421 TTCTCGCGCAACGGGTGATGCGGTGATCAATTAATATCCGCTGATGACACAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCGCTGCACTAATGTTCCGGCGTATTTCTTGATGCTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCGCTGCACTAATGTTCCGGCGTATTTCTTGATGCTCTGACCCAGA 540
QY 541 CACCCATCAACAGATATTAATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGATATTAATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
QY 601 TGGTGGATTTGGGTACACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTCTGTCGG 660
Db 601 TGGTGGATTTGGGTACACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTCTGTCGG 660
QY 661 CGCGTCTCGGCTCGGCTGGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGGCTCGGCTGGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCTGTTCCCATCTGCGATGCTGGTTCGCAACGATCAGATGCGCTGGGGCGAA 840
Db 781 ATGAGGGCATCTGTTCCCATCTGCGATGCTGGTTCGCAACGATCAGATGCGCTGGGGCGAA 840
QY 841 TGGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAGC 900
Db 841 TGGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAGC 900
QY 901 AGGATACCGAAGACAGCTCATGTTATATCCGCGCTTAACCAACCATCAAAACAGGATTTTC 960
Db 901 AGGATACCGAAGACAGCTCATGTTATATCCGCGCTTAACCAACCATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCTCTCACTGTTGAAAAGAAAACCACTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCCTCTCACTGTTGAAAAGAAAACCACTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTAATGAGCTGGCGACGACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTAATGAGCTGGCGACGACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTGTAGCTCACTATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTGTAGCTCACTATTAG 1200
QY 1201 GCACAATTTCTATGTTTGACAGCTTATCATCGACTGCGGTGACCAATGCTTCTGGCG 1260
```

Db 1201 GCACAAATTCATGTTTGACAGCTTATCATCACTGACGGTGCCACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTCGTAAATCACTGCATATTCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTCGTAAATCACTGCATATTCG 1320
Qy 1321 TGTGCGTCAAGGCGCATCTCCGTTCTGGATAATGTTTTTGGCGCGACATCATACGGTT 1380
Db 1321 TGTGCGTCAAGGCGCATCTCCGTTCTGGATAATGTTTTTGGCGCGACATCATACGGTT 1380
Qy 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
Db 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
Qy 1441 ATTGTGAGCGGATAACAAATTTTACACAGGAAACAGCAGTCGGTTAGTGTTTTTACGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTTTACACAGGAAACAGCAGTCGGTTAGTGTTTTTACGA 1500
Qy 1501 GCACCTTCAACAAAGGACCATAGATTATGAAAACTGAAGAGGTAATCTGGTAATCTGG 1560
Db 1501 GCACCTTCAACAAAGGACCATAGATTATGAAAACTGAAGAGGTAATCTGGTAATCTGG 1560
Qy 1561 ATTAACGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGTT 1680
Qy 1681 GCGCAATCGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAATCGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGCTGAATACACCCCGGACAAAGGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAATACACCCCGGACAAAGGTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGATGCGGTACGTTTACACGGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGGTACGTTTACACGGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
Qy 1861 GAAGGTTATCGCTGATTATTAACAAGATCTGTGCGGAACCCGCGCAAAACCTCGGAA 1920
Db 1861 GAAGGTTATCGCTGATTATTAACAAGATCTGTGCGGAACCCGCGCAAAACCTCGGAA 1920
Qy 1921 GAGATCCCGGCTGGATAAAGAACTGAAGCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCTGGATAAAGAACTGAAGCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGTCGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGTCGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAACGCGAAGTACGACATTAAGACGTTGGCGTGGATACGCTGGCGGAAGCG 2100
Db 2041 TATGAAACGCGAAGTACGACATTAAGACGTTGGCGTGGATACGCTGGCGGAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGTTGTCCTGATTAAACAAACACATGAATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGTCCTGATTAAACAAACACATGAATGACAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGNAGTGGCTTTTAAGAGGCGAAACAGCGATGACCATCAACGCGCCGTGG 2220
Db 2161 TCCATCGCAGNAGTGGCTTTTAAGAGGCGAAACAGCGATGACCATCAACGCGCCGTGG 2220
Qy 2221 GCATGGTCCAACTCGACACAGCAAAATGAAATTTATGTTGTAAGGTAAGTCCGACCTTC 2280
Db 2221 GCATGGTCCAACTCGACACAGCAAAATGAAATTTATGTTGTAAGGTAAGTCCGACCTTC 2280
Qy 2281 AAGGGTCAACCATCCAAACCGTTCGTTGGCTGCTGAGCGCAGGTATTAACGCGCCGAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTCGTTGGCTGCTGAGCGCAGGTATTAACGCGCCGAGT 2340

RESULT 4

US-09-765-555-17

; Sequence 17, Application US/09765555

; Publication No. US20030037355A1

; GENERAL INFORMATION:

; APPLICANT: The Scripps Research Institute

; TITLE OF INVENTION: Methods and compositions to modulate

Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACTATCTGCTGACTGATGAAGGTCGT 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACTATCTGCTGACTGATGAAGGTCGT 2400
Qy 2401 GAAGCGGTTAATAAAGACAAACCCCTGGGTGCGGTAGCGTGAAGTCTTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCCCTGGGTGCGGTAGCGTGAAGTCTTTACGAGGAAG 2460
Qy 2461 TTGCGGAAAAGATCCAGTATTTGCGGCCACCATGGAACCCCAAGAAAGGTGAAATCATG 2520
Db 2461 TTGCGGAAAAGATCCAGTATTTGCGGCCACCATGGAACCCCAAGAAAGGTGAAATCATG 2520
Qy 2521 CCGAACATCCGCGAGATGTCGGTTCCTGTTGTTGTCGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCGCGAGATGTCGGTTCCTGTTGTTGTTGTCGCTGCTGCTGCTGCTGCTG 2580
Qy 2581 GCCAGCGGTGCTCAGACTGTCGATGAAGCCTGAAAGACGCGAGACTTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGCTCAGACTGTCGATGAAGCCTGAAAGACGCGAGACTTAATTCGAGCTCG 2640
Qy 2641 AACCAACAAACAAATAAACAACAACCTCGGGATCGAGGAAAGGATTTTCAGAAATTC 2700
Db 2641 AACCAACAAACAAATAAACAACAACCTCGGGATCGAGGAAAGGATTTTCAGAAATTC 2700
Qy 2701 GGATCTCTTCTGTCGCGCCAGGGGCGCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTGTCGCGCCAGGGGCGCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
Qy 2761 CCGGAATGTTGTAAGTCTTCTCTCAGAGCTCTCACCTGCTGGGCCACACAGCGTACCCAC 2820
Db 2761 CCGGAATGTTGTAAGTCTTCTCAGAGCTCTCACCTGCTGGGCCACACAGCGTACCCAC 2820
Qy 2821 ACGGGTGAAAAACCGGTATAATGCCAGAGTGGCGCAAAATCTTTTACGCAAGTCCAGCAAC 2880
Db 2821 ACGGGTGAAAAACCGGTATAATGCCAGAGTGGCGCAAAATCTTTTACGCAAGTCCAGCAAC 2880
Qy 2881 CTGCTGGCCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCAGAAATGTTGGC 2940
Db 2881 CTGCTGGCCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCAGAAATGTTGGC 2940
Qy 2941 AAGTCTTCTCTCGGTCTGACAAATCTCTGTCGCGCACCAAGTACTCACACCGGGGAGAAG 3000
Db 2941 AAGTCTTCTCAGCCAGTCCAGCAACCTGGTGGCGCCACCAAGTACTCACACCGGGGAGAAG 3000
Qy 3001 CCCTATGCTGTCGCGAATGTTGTAAGTCTTCTCAGCGCGAGCGATTAACCTGTTGCGCCAC 3060
Db 3001 CCCTATGCTGTCGCGAATGTTGTAAGTCTTCTCAGCAACCTGGTGGCTCTCTTGGTTAGCAC 3060
Qy 3061 CAGCGTACCCACACGCGGTGAAAAACCGGTATAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACGCGGTGAAAAACCGGTATAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
Qy 3121 CAGCGCGCCACCTGGAACGCCATCAACGCACTCATACTGGCGAGAAGCCATACAAATGT 3180
Db 3121 CAGCGCGCCACCTGGAACGCCATCAACGCACTCATACTGGCGAGAAGCCATACAAATGT 3180
Qy 3181 CCAGAATGTTGGCAAGTCTTCTCTCGTCTGCAATCTCTGTCGCGGACCAACGTAAGTCTAC 3240
Db 3181 CCAGAATGTTGGCAAGTCTTCTCTCAACTTCAAGTCTTGGTCCGTCAACCAAGTCTAC 3240
Qy 3241 ACCGGTAAAAAAACTAGTGGCGCGCGCCAGTACCCGTACGACGTTCCGGACTACGCT 3300
Db 3241 ACCGGTAAAAAAACTAGTGGCGCGCGCCAGTACCCGTACGACGTTCCGGACTACGCT 3300

; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m4 and zinc finger
; OTHER INFORMATION: protein 2Ppm4
US-09-765-555-17

Query Match 97.4%; Score 3215.2; DB 10; Length 3300;

Best Local Similarity 98.4%; Pred. No. 0;

Matches 3247; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY	1	CGGACACCATCGAATGGTGC	AAACCTTTCCGGGTATG	GCATGATAGCGCCGGAAGAGA	60
DB	1	CGGACACCATCGAATGGTGC	AAACCTTTCCGGGTATG	GCATGATAGCGCCGGAAGAGA	60
QY	61	GTCATTCAGGGTGGTGAATG	TGAACACAGTAACCGT	TATACGATGTCGACAGTATGCGG	120
DB	61	GTCATTCAGGGTGGTGAATG	TGAACACAGTAACCGT	TATACGATGTCGACAGTATGCGG	120
QY	121	GTGTCCTTATCAGACCGT	TTCCCGGTGGAACACG	AGCCAGCCACGCTTTCTCGAAAA	180
DB	121	GTGTCCTTATCAGACCGT	TTCCCGGTGGAACACG	AGCCAGCCACGCTTTCTCGAAAA	180
QY	181	CGCGGAAAAAGTGAAGCGG	CGATGGCGGAGCTGAAT	TACATTTCCACCCGCTGGCAC	240
DB	181	CGCGGAAAAAGTGAAGCGG	CGATGGCGGAGCTGAAT	TACATTTCCACCCGCTGGCAC	240
QY	241	AACAACTGCGGGCAAA	CAGTCTGTTGCTGAT	TGGCGTTGCCACCTTCCAGTCTGGCCCTGC	300
DB	241	AACAACTGCGGGCAAA	CAGTCTGTTGCTGAT	TGGCGTTGCCACCTTCCAGTCTGGCCCTGC	300
QY	301	ACGCGCGCTCGCAAA	TGTTCGGCGGATTAAT	CTCGCGCGATCAACTGGGTGCCAGCG	360
DB	301	ACGCGCGCTCGCAAA	TGTTCGGCGGATTAAT	CTCGCGCGATCAACTGGGTGCCAGCG	360
QY	361	TGGTGTGTCGATGTA	GAACAAAGGCGCTCGA	AGCCCTGTAAAGCGGCGTGCACAATC	420
DB	361	TGGTGTGTCGATGTA	GAACAAAGGCGCTCGA	AGCCCTGTAAAGCGGCGTGCACAATC	420
QY	421	TTCTCGCGCAACGCT	CAGTGGGCTGATCAT	TAACTATTCGCTGGATGACCAAGGATGCCA	480
DB	421	TTCTCGCGCAACGCT	CAGTGGGCTGATCAT	TAACTATTCGCTGGATGACCAAGGATGCCA	480
QY	481	TTGCTGTGGAAGCT	GCCTGCCTCACTAAT	TGTTCCGCGGTTATTTCTTGATGCTCTGACCCAGA	540
DB	481	TTGCTGTGGAAGCT	GCCTGCCTCACTAAT	TGTTCCGCGGTTATTTCTTGATGCTCTGACCCAGA	540
QY	541	CACCCATCAACAGTA	TATTTTCTCCATGA	AGACGGTACCGGCTGGGCGTGGAGCATC	600
DB	541	CACCCATCAACAGTA	TATTTTCTCCATGA	AGACGGTACCGGCTGGGCGTGGAGCATC	600
QY	601	TGGTTCGATTTGGGT	CACAGCAAACTCGC	CTGTTAGCGGGCCCATTAAGTCTGTCCTGG	660
DB	601	TGGTTCGATTTGGGT	CACAGCAAACTCGC	CTGTTAGCGGGCCCATTAAGTCTGTCCTGG	660
QY	661	CGCGTCTGGCTG	CTGGCTGGCATAAAT	TCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	661	CGCGTCTGGCTG	CTGGCTGGCATAAAT	TCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAACGGGAAGCG	CACTGGAGTGCAT	TCGCGTTTCAACAAACCATGCAAAATGCTGA	780
DB	721	CGGAACGGGAAGCG	CACTGGAGTGCAT	TCGCGTTTCAACAAACCATGCAAAATGCTGA	780

DB	721	CGGAACGGGAAGCGCA	CTGGAGTGCCAT	TGTCGGGTTTCAACAAACCATGCAAAATGCTGA	780
QY	781	ATGAGGGCATCGTTCC	CACTGCGATGCT	TGGTGGCAACGATCAGATGCGCTGGCGCGAA	840
DB	781	ATGAGGGCATCGTTCC	CACTGCGATGCT	TGGTGGCAACGATCAGATGCGCTGGCGCGAA	840
QY	841	TGGCGGCCATTACCG	AGTCCGGCTCGCG	TGTTGGTGGCGGATATCTCGGTAGTGGGATAGG	900
DB	841	TGGCGGCCATTACCG	AGTCCGGCTCGCG	TGTTGGTGGCGGATATCTCGGTAGTGGGATAGG	900
QY	901	ACGATACCGAAGACA	GAGCTCATGTTAT	TATCCCGCGCTTAACCAACCATCAAAACGAGATTTTC	960
DB	901	ACGATACCGAAGACA	GAGCTCATGTTAT	TATCCCGCGCTTAACCAACCATCAAAACGAGATTTTC	960
QY	961	GCCTGCTGGGCAAA	ACCAAGCGTGGAC	CGCTTCTGCAACTCTCTCAGGCGCCAGGCGTGA	1020
DB	961	GCCTGCTGGGCAAA	ACCAAGCGTGGAC	CGCTTCTGCAACTCTCTCAGGCGCCAGGCGTGA	1020
QY	1021	AGGGCAATCAGCT	GTGGCCCTCTCA	CTGCTGTAAGAAAGAAACCAACCTTGGCGCCCAATA	1080
DB	1021	AGGGCAATCAGCT	GTGGCCCTCTCA	CTGCTGTAAGAAAGAAACCAACCTTGGCGCCCAATA	1080
QY	1081	CGCAAAACCGCT	CTCCCGCGCTT	GGCGGATTCATTAAATGCAGCTGGCAGACAGGTTT	1140
DB	1081	CGCAAAACCGCT	CTCCCGCGCTT	GGCGGATTCATTAAATGCAGCTGGCAGACAGGTTT	1140
QY	1141	CCGACTGGAAGCGG	GCAGTGAGCGCA	ACCAATTAATGTGAGTTAGCTCACTCATTTAG	1200
DB	1141	CCGACTGGAAGCGG	GCAGTGAGCGCA	ACCAATTAATGTGAGTTAGCTCACTCATTTAG	1200
QY	1201	GCACAAATTCAT	TTTGACAGCTTAT	CATCGCTGACCGTGACCAATGCTTCTGGG	1260
DB	1201	GCACAAATTCAT	TTTGACAGCTTAT	CATCGCTGACCGTGACCAATGCTTCTGGG	1260
QY	1261	TCAGCAGCCAT	CGGAAGCTGT	GTGATGCTGTCAGCTGTAATCACTGCATAAATTCG	1320
DB	1261	TCAGCAGCCAT	CGGAAGCTGT	GTGATGCTGTCAGCTGTAATCACTGCATAAATTCG	1320
QY	1321	TGTCGCTCAAGCG	CACTCCCGTTCT	TGGATAATGTTTTTGGCGCGACATATAACGGTT	1380
DB	1321	TGTCGCTCAAGCG	CACTCCCGTTCT	TGGATAATGTTTTTGGCGCGACATATAACGGTT	1380
QY	1381	CTGGCAAAATAT	TTTGAATGAGCT	TGTTGACAAATTAATCATCGGCTCGTATTAATGTGTGA	1440
DB	1381	CTGGCAAAATAT	TTTGAATGAGCT	TGTTGACAAATTAATCATCGGCTCGTATTAATGTGTGA	1440
QY	1441	ATTGTGAGCGGAT	AACAAATTTCA	CACAGGAACCGCAGTCCGTTAGTGTTCACGA	1500
DB	1441	ATTGTGAGCGGAT	AACAAATTTCA	CACAGGAACCGCAGTCCGTTAGTGTTCACGA	1500
QY	1501	GCACCTTCAACCA	AGGACCATAGAT	TATGAAACTGGAAGAGTAAATCTGGTAATCTGG	1560
DB	1501	GCACCTTCAACCA	AGGACCATAGAT	TATGAAACTGGAAGAGTAAATCTGGTAATCTGG	1560
QY	1561	ATTAAACCGCGAT	AAAGGCTATAA	CGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
DB	1561	ATTAAACCGCGAT	AAAGGCTATAA	CGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAANG	TACCGTTGAG	CATCCGATTAACCTGGAAGAGAAATTCACAGGTT	1680
DB	1621	ACCGGAATTAANG	TACCGTTGAG	CATCCGATTAACCTGGAAGAGAAATTCACAGGTT	1680
QY	1681	CGGCAACTGCGAT	TGGCCCTGAC	ATTATCTTGGGCAACAGCCGCTTTGGTGGCTAC	1740
DB	1681	CGGCAACTGCGAT	TGGCCCTGAC	ATTATCTTGGGCAACAGCCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATTCGCT	TGTTGGCTGAA	ATCAACCCGGCAAAAGCGTTCCAGGAACAAGCTGAT	1800
DB	1741	GCTCAATTCGCT	TGTTGGCTGAA	ATCAACCCGGCAAAAGCGTTCCAGGAACAAGCTGAT	1800
QY	1801	CGTTTACCTGG	ATGCGTACG	TTTCAACCGCAAGCTGATGCTTACCCGATCCCTGTT	1860
DB	1801	CGTTTACCTGG	ATGCGTACG	TTTCAACCGCAAGCTGATGCTTACCCGATCCCTGTT	1860

Qy	1961	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCGCGCAAAAAA	CTGGGAA	1921
Db	1961	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCGCGCAAAAAA	CTGGGAA	1920
Qy	1921	GAGATCCGGCGCTGGATAAAGAACTGAAGCGAAAGTGAAGACGCGCTGATGTTCAAC	1980	
Db	1921	GAGATCCGGCGCTGGATAAAGAACTGAAGCGAAAGTGAAGACGCGCTGATGTTCAAC	1980	
Qy	1981	CTGCAAGAACGGTACTTCACTGCGCGCTGATGCTCTGACGGGGTTATGCTTCAAG	2040	
Db	1981	CTGCAAGAACGGTACTTCACTGCGCGCTGATGCTCTGACGGGGTTATGCTTCAAG	2040	
Qy	2041	TATGAAACCGCAAGTACGACATTAAGACGTTGGCGTGTGATACGCTGCGCGAAGCG	2100	
Db	2041	TATGAAACCGCAAGTACGACATTAAGACGTTGGCGTGTGATACGCTGCGCGAAGCG	2100	
Qy	2101	GGTCTGACCTTCTCTGGTTGACCTGATTAAAAAACAACATGATGACAGACACCGATTAC	2160	
Db	2101	GGTCTGACCTTCTCTGGTTGACCTGATTAAAAAACAACATGATGACAGACACCGATTAC	2160	
Qy	2161	TCCATCCGAGAAGTGTCTTTAATAAAGCGAAACAGCGATGACCATCAACGCCCGTGG	2220	
Db	2161	TCCATCCGAGAAGTGTCTTTAATAAAGCGAAACAGCGATGACCATCAACGCCCGTGG	2220	
Qy	2221	GCATGTTCCACATCGACACACGCAAGTGAATTAATGGTGTACGGTACTGCGCACTTC	2280	
Db	2221	GCATGTTCCACATCGACACACGCAAGTGAATTAATGGTGTACGGTACTGCGCACTTC	2280	
Qy	2281	AAGGGTCAACCATCCAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT	2340	
Db	2281	AAGGGTCAACCATCCAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT	2340	
Qy	2341	CCGAAACAAAGAGTGGCAAAAGAGTTCCTCGAAAACTATCTGTGACTGATGAAGGTCTG	2400	
Db	2341	CCGAAACAAAGAGTGGCAAAAGAGTTCCTCGAAAACTATCTGTGACTGATGAAGGTCTG	2400	
Qy	2401	GAAGCGTTAATAAGACAAACCGCTGGTGGCGTACGCTGAAGTCTTACGAGGAGAG	2460	
Db	2401	GAAGCGTTAATAAGACAAACCGCTGGTGGCGTACGCTGAAGTCTTACGAGGAGAG	2460	
Qy	2461	TTGGCGAAAGATCCAAGTATTTGGCGCCACCATGGAAAAACCGCCAGAAAGTGAATCATG	2520	
Db	2461	TTGGCGAAAGATCCAAGTATTTGGCGCCACCATGGAAAAACCGCCAGAAAGTGAATCATG	2520	
Qy	2521	CCGAAACATCCCGCAGATGTCGGTTCCTGATGTCGGTGTGCGTACTGCGTGAACGCC	2580	
Db	2521	CCGAAACATCCCGCAGATGTCGGTTCCTGATGTCGGTGTGCGTACTGCGTGAACGCC	2580	
Qy	2581	GCCAGCGGTCTGAGACTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640	
Db	2581	GCCAGCGGTCTGAGACTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640	
Qy	2641	AACAACAAACAATAACAAATAACAAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC	2700	
Db	2641	AACAACAAACAATAACAAATAACAAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC	2700	
Qy	2701	GGATCTCTTCTCTGTGGCCAGGCGCCCTCGAGCCCGGGGAGAGCCCTATGCTTGT	2760	
Db	2701	GGATCTCTTCTCTGTGGCCAGGCGCCCTCGAGCCCGGGGAGAGCCCTATGCTTGT	2760	
Qy	2761	CCGGAATGTGGTAAGTCTCTCTCAGAGCTCTCAGCTGTGGCCACACGAGCTACCCAC	2820	
Db	2761	CCGGAATGTGGTAAGTCTCTCTCAGAGCTCTCAGCTGTGGCCACACGAGCTACCCAC	2820	
Qy	2821	ACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTATGCCAGTCCAGCAAC	2880	
Db	2821	ACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTATGCCAGTCCAGCAAC	2880	
Qy	2881	CTGGTGGCCATCAACGCACTCATCTGCGGAGAGGCCATACAAATGTCCAGAAATGTGGC	2940	
Db	2881	CTGGTGGCCATCAACGCACTCATCTGCGGAGAGGCCATACAAATGTCCAGAAATGTGGC	2940	

QY	2941	AAGTCTTTCTTCGGTCTGCAAAATCTCGTCCGGCACCAAAGTACTCACAACGGGGGGAAG	3000
Db	2941	AAGTCTTTCTTCAGTGAATTCGTGTGATCTTTGCGAGGCACCAAAGTACTCACAACGGGGGGAAG	3000
QY	3001	CCCTATGCTTGTCCGGAAATGGTAAAGTCTTTCAGCCGCAAGCATACCTGGTGGCGGCAC	3060
Db	3001	CCCTATGCTTGTCCGGAAATGGTAAAGTCTTCTCTCAGAGCTCTCACCTGGTGGCGGCAC	3060
QY	3061	CAGCGTACCCACACCGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGC	3120
Db	3061	CAGCGTACCCACACCGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGC	3120
QY	3121	CAGCCGGCCACCTGGCCAGCCATCAACGCACTCAATCTCTCGGAGAGCCATACAAATGT	3180
Db	3121	CGCAGCGATTAACCTGGTGGCCCATCAACGCACTCATATCTGCGAGAGAGCCATACAAATGT	3180
QY	3181	CCAGAAATGTGGCAAGTCTTTCTCTCGGTCTGCAAAATCTCGTCCGGCACCACAGTACTCAC	3240
Db	3181	CCAGAAATGTGGCAAGTCTTTCTCNACTTCAGGCCATTTGGTCCGTCAACCAAGTACTCAC	3240
QY	3241	ACCGGTAAAAAACTAGTGGCCAGGCGCGGCGCAGTACCCTGACGACGTTCGGGACTACGCT	3300
Db	3241	ACCGGTAAAAAACTAGTGGCCAGGCGCGGCGCAGTACCCTGACGACGTTCGGGACTACGCT	3300

RESULT 5

US-09-765-555-16

; Sequence 16, Application US/09765555

; Publication No. US20030037355A1

GENERAL INFORMATION:

; APPLICANT: The Scripps Research Institute

; TITLE OF INVENTION: Methods and compositions to modulate

; TITLE OF INVENTION: expression in plants

; FILE REFERENCE: 27801-20014.40

; CURRENT APPLICATION NUMBER: US/09/765,555

; CURRENT FILING DATE: 2002-05-24

; PRIOR APPLICATION NUMBER: US 09/620,897

; PRIOR FILING DATE: 2000-01-21

; PRIOR APPLICATION NUMBER: US 60/177,468

; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 75

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 16

; LENGTH: 3300

; TYPE: DNA

; ORGANISM: Artificial Sequence

FEATURE:

; OTHER INFORMATION: Partial sequence of pMal-m3 and zinc finger

; OTHER INFORMATION: protein ZFPm3

US-09-765-555-16

Query Match	97.2%	Score 3207.2	DB 10	Length 3300
Best Local Similarity	98.2%	Pred. No. 0		
Matches 3242	Conservative 0	Mismatches 58	Indels 0	Gaps 0
Qy	1	CCGACACCATCGAATGGTGTCAAAACCTTTCCGCGGTATGGCATGATAGCGCCCGAAGAGA	60	
Db	1	CCGACACCATCGAATGGTGTCAAAACCTTTCCGCGGTATGGCATGATAGCGCCCGAAGAGA	60	
Qy	61	GTCAATTACGGTGGTGAATTGTCAAAACCAAGTAACGTTTATACGATGTTCGACAGATATGCCG	120	
Db	61	GTCAATTACGGTGGTGAATTGTCAAAACCAAGTAACGTTTATACGATGTTCGACAGATATGCCG	120	
Qy	121	GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACAGCGCAGCCACCGTTCCTTCGCAAAA	180	
Db	121	GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACAGCGCAGCCACCGTTCCTTCGCAAAA	180	
Qy	181	CGCGGGAAAAAGTGGAAAGCGCGCATGGCGAGCTGAATTACATTCCCAACCGCGTGGCAC	240	
Db	181	CGCGGGAAAAAGTGGAAAGCGCGCATGGCGAGCTGAATTACATTCCCAACCGCGTGGCAC	240	
Qy	241	AACCAACTGGCGGGCAACACGTCGTGTGATTTGGCGTTTGGCACCTCCAGTCTCGGCCCTGC	300	

Db 241 AACAACTGGCGGCAAA CAGTGGTCTGATGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGCGCCGTCGCAAA TGTGCGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGC 360
Db 301 ACGCGCCGTCGCAAA TGTGCGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGC 360
Qy 361 TGGTGGTCTGATGTA GAAAGAGCGCGTGAAGCGCTGAAAGCGCGGTGCAACAATC 420
Db 361 TGGTGGTCTGATGTA GAAAGAGCGCGTGAAGCGCTGAAAGCGCGGTGCAACAATC 420
Qy 421 TTCTCGCCCAACGCTC AGTGGGCTGATCATTAATCTCGCTGATACACGAGATGCCA 480
Db 421 TTCTCGCCCAACGCTC AGTGGGCTGATCATTAATCTCGCTGATACACGAGATGCCA 480
Qy 481 TTGCTGTGGAAGCTC GCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTCCTGACAGA 540
Db 481 TTGCTGTGGAAGCTC GCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTCCTGACAGA 540
Qy 541 CACCCATCAACAGTAT TATTTCTCCCATGAAGCGGTACGGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTAT TATTTCTCCCATGAAGCGGTACGGACTGGCGGTGGAGCATC 600
Qy 601 TGGTGGCATGGGTCA CAGCAAAATCGCGCTGTTAGCGGCCCATTTAAGTTCTGTCCTCG 660
Db 601 TGGTGGCATGGGTCA CAGCAAAATCGCGCTGTTAGCGGCCCATTTAAGTTCTGTCCTCG 660
Qy 661 CGCGTCTCGCTGCTG GCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCGATAG 720
Db 661 CGCGTCTCGCTGCTG GCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGCA CTGGAGTGCATGTCGGTTTTCAACAAACCATGCAATGCTGA 780
Db 721 CGGAACGGGAAGGCA CTGGAGTGCATGTCGGTTTTCAACAAACCATGCAATGCTGA 780
Qy 781 ATGAGGGCATGCTTCC CACTGCGATGCTGGTGGCCAAACGATCAGATGGCGTGGCGCAA 840
Db 781 ATGAGGGCATGCTTCC CACTGCGATGCTGGTGGCCAAACGATCAGATGGCGTGGCGCAA 840
Qy 841 TGGCGCCATTACCAAG TCCGGGTCGCGGTGGTGGCGATATCTCGTAGTGGGATACG 900
Db 841 TGGCGCCATTACCAAG TCCGGGTCGCGGTGGTGGCGATATCTCGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACG TCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACG TCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
Qy 961 GCCTGCTGGGGCAAC CAGCGTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGGCAAC CAGCGTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGT TGGCCGCTCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGT TGGCCGCTCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCCTCTCC CCGCGGTTGGCGATTCATTAATGAGCTGGCAACAGGTTT 1140
Db 1081 CGCAACCGCCTCTCC CCGCGGTTGGCGATTCATTAATGAGCTGGCAACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGG CAGTGACGCAACGCAATTAATGTCAGTTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGG CAGTGACGCAATTAATGTCAGTTAGTCACTCATTTAG 1200
Qy 1201 GCACAATCTCATGTT TGAACGCTTATCATCGACTGCAGGTGCAACCAATGCTCTGGGG 1260
Db 1201 GCACAATCTCATGTT TGAACGCTTATCATCGACTGCAGGTGCAACCAATGCTCTGGGG 1260
Qy 1261 TCAGGAGCCATCGGA AGCTGTGGTATGAGTGTGAGCTGCGAGTTCGTAATCACTGCAATTCG 1320
Db 1261 TCAGGAGCCATCGGA AGCTGTGGTATGAGTGTGAGCTGCGAGTTCGTAATCACTGCAATTCG 1320
Qy 1321 TGTTCCTCAAGCGCA CTCCCGCTCTGGATATGTTTTTGGCGGACATCATACGGTT 1380
Db 1321 TGTTCCTCAAGCGCA CTCCCGCTCTGGATATGTTTTTGGCGGACATCATACGGTT 1380

Qy 1381 CTGCGCAAAATATCTG AAATGAGCTGTTGA CAATTAATCATCGGCTCGTATTAATGTGTGGA 1440
Db 1381 CTGCGCAAAATATCTG AAATGAGCTGTTGA CAATTAATCATCGGCTCGTATTAATGTGTGGA 1440
Qy 1441 ATTGTGAGCGGATAA CAAATTTTACACAGGAAACCGAGTCGTTTAGTGTGTTTTCAGCA 1500
Db 1441 ATTGTGAGCGGATAA CAAATTTTACACAGGAAACCGAGTCGTTTAGTGTGTTTTCAGCA 1500
Qy 1501 GCATTTCAACCAAGGA CCAATAGATTATGAAAA CTGAAGAGGTAAACTGTTAACTCTGG 1560
Db 1501 GCATTTCAACCAAGGA CCAATAGATTATGAAAA CTGAAGAGGTAAACTGTTAACTCTGG 1560
Qy 1561 ATTAACGGCGATAAAG GCTATAACCGTCTCGCTGAAGTCGTTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAG GCTATAACCGTCTCGCTGAAGTCGTTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCA CCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCCTCCACAGTT 1680
Db 1621 ACCGGAATTAAGTCA CCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCCTCCACAGTT 1680
Qy 1681 CGCGCAACTGGCGATG GCGCCTTGACATTTCTTCTGGGCACACGACCGCTTTTGGTGGCTAC 1740
Db 1681 CGCGCAACTGGCGATG GCGCCTTGACATTTCTTCTGGGCACACGACCGCTTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCTGT TGGCTGAAATCAACCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGT TGGCTGAAATCAACCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATG CCGTACGTTTACAACCGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATG CCGTACGTTTACAACCGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
Qy 1861 GAAGCGTTATCGCTGA TTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGA TTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGG ATAAAGAACTGAAACGGAAAGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGG ATAAAGAACTGAAACGGAAAGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACT TTTACCTGCGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACT TTTACCTGCGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAAACGCAAGT ACAGATTAAGACCGTGGCGCTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGT ACAGATTAAGACCGTGGCGCTGGATTAACGCTGGCGCGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTG GTTGAACCTGATTTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTG GTTGAACCTGATTTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGAGAGCTGCT TTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGAGAGCTGCT TTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCCAAATCGA CACCCAGCAAAAGTGAATTTATGGTGAACCGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGA CACCCAGCAAAAGTGAATTTATGGTGAACCGTACTGCCGACCTTC 2280
Qy 2281 AAGGTCACCATCAAA ACCGTTCTGTTGGCTGCTGAGCGAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTCACCATCAAA ACCGTTCTGTTGGCTGCTGAGCGAGGTATTAACGCCGCCAGT 2340
Qy 2341 CCGAACAAGAGCTGGC AAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
Db 2341 CCGAACAAGAGCTGGC AAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
Qy 2401 GAAGCGGTTAATAAGA CAAACCGCTGGGTGCTGAGCGCTGAACTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGA CAAACCGCTGGGTGCTGAGCGCTGAACTCTTACGAGGAAG 2460


```
QY 2461 TTGGCGAAAGATCCAGTATTGCGCCACCACCATGGAAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTGCGCCACCACCATGGAAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CGGAACATCCCGCAGATGTCGGCTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CGGAACATCCCGCAGATGTCGGCTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGTCAGATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGTCAGATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAACAATAACAACCACTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACCACTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
QY 2701 GGATCCTCTTCTCTGTCGGCCAGCGGCCCTCGAGCCCGGGGAGAAGCCCTATGCTTGT 2760
Db 2701 GGATCCTCTTCTCTGTCGGCCAGCGGCCCTCGAGCCCGGGGAGAAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTGTAAGTCTTCTCTCAGAGCTCTCACCTGGTGGCCACCCAGCGTACCCAC 2820
Db 2761 CCGGAATGTGTAAGTCTTCTCAGAGCTCTCACCTGGTGGCCACCCAGCGTACCCAC 2820
QY 2821 ACGGTGAAACCGTATATAATGCCAGAGTCCGCAAACTCTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGTGAAACCGTATATAATGCCAGAGTCCGCAAACTCTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGGTGGCCATCAACGCACTCATACTGGCGAGAAGCCATACAAATGTCAGAAATGTGGC 2940
Db 2881 CTGGTGGCCATCAACGCACTCATACTGGCGAGAAGCCATACAAATGTCAGAAATGTGGC 2940
QY 2941 AAGTCTTTCTCGGTCTGACAAATCTGTCGGCACCAACGTACTCTCACCCGGGAGAAG 3000
Db 2941 AAGTCTTTCTCGGTCTGACAAATCTGTCGGCACCAACGTACTCTCACCCGGGAGAAG 3000
QY 3001 CCCTATGCTGTCGGGAATGTTAAGTCTTCAAGCTTCAAGCCGAGGATACTGGTGGCCAC 3060
Db 3001 CCCTATGCTGTCGGGAATGTTAAGTCTTCAAGCTTCTTCAAGCCGAGGATACTGGTGGCCAC 3060
QY 3061 CAGCGTACCCACAGCGGTGAAAAACCGTATATAATGCGCAGAGTCCGCAAACTCTTTTAGC 3120
Db 3061 CAGCGTACCCACAGCGGTGAAAAACCGTATATAATGCGCAGAGTCCGCAAACTCTTTTAGC 3120
QY 3121 CAGCGCGGCACCTGGCCAGCCATCAACGCACTCATCTGCGGAGAAGCCATACAAATGT 3180
Db 3121 GACTGCGCGCAGCTTGTCTGCGCCATCAACGCACTCATCTGCGGAGAAGCCATACAAATGT 3180
QY 3181 CCAGAAATGTGGCAAGTCTTCTCTCGGTCTGACAAATCTGTCGGCACCAACGTAACCTAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTCTCTCGGTCTGACAAATCTGTCGGCACCAACGTAACCTAC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGGCGGAGTACCGTACGACGTTCCGGGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGGCGGAGTACCGTACGACGTTCCGGGACTACGCT 3300

RESULT 6
US-10-257-384A-1
; Sequence 1, Application US/10257384A
; Publication No. US20040087524A1
; GENERAL INFORMATION:
; APPLICANT: Wiedersanders, Bernd
; APPLICANT: Maubach, Gunter
; TITLE OF INVENTION: Agent for postoperative use after removal of bone tumors
; FILE REFERENCE: 2945-101
; CURRENT APPLICATION NUMBER: US/10/257,384A
; PRIOR FILING DATE: 2003-08-21
; PRIOR APPLICATION NUMBER: PCT/DE 01/01510
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: DE 100 20 125.3
; PRIOR FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 19
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 7373
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Expression vector pMalc2cchbmp2
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)..(3453)
; OTHER INFORMATION: /product = fusion protein MBP Cystatin C BMP-2
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2689)
; OTHER INFORMATION: Faktor Xa Protease /position=388
; OTHER INFORMATION: (Amino acid sequence)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3067)..(3108)
; OTHER INFORMATION: /spacer between Cystatin C andBMP-2
; PUBLICATION INFORMATION:
; AUTHORS: Zwick, M B.
; JOURNAL: Anal. Biochem.
; VOLUME: 264
; ISSUE: 1
; PAGES: 87-97
; DATE: 1998-11-01
; DATABASE ACCESSION NUMBER: Entrez Nucleotide database/AF031813
; DATABASE ENTRY DATE: 2001-05-07
; RELEVANT RESIDUES: 1528 - 3453
US-10-257-384A-1
```

```
Query Match 82.0%; Score 2705.8; DB 18; Length 7373;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2719; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAAAACCTTTCGGGGTATGCATGATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTCGGGGTATGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTCAGGGTGGTGAATGTGAACACCAAGTAACTATACGATGTCGAGAGTATGCCG 120
Db 61 GTCAATTCAGGGTGGTGAATGTGAACACCAAGTAACTATACGATGTCGAGAGTATGCCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCCGGTGTGAACCCAGCCAGCCACGTTTCTGCCAAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCCGGTGTGAACCCAGCCAGCCACGTTTCTGCCAAAA 180
QY 181 CGCGGAAAAAGTGGAAAGCGCGGATGCGCGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGGAAAGCGCGGATGCGCGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGCAACAGTGTGCTGATGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAACAGTGTGCTGATGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGCTCGCAAAATTTGCGCGCGGATTAATTTCTCGGCCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGCTCGCAAAATTTGCGCGCGGATTAATTTCTCGGCCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGTGTGCGATGTGTAGAACGAAAGCGCGTCTGAAAGCGCGGTGTAAGCGGGTGCAATC 420
Db 361 TGGTGTGTGCGATGTGTAGAACGAAAGCGCGTCTGAAAGCGCGGTGTAAGCGGGTGCAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGTGATCAATTAATTCCTCCGCTGGATGACAGAGTGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGTGATCAATTAATTCCTCCGCTGGATGACAGAGTGCCA 480
QY 481 TTGCTGTGGAAGCTGCCTGCATTAATTTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
Db 481 TTGCTGTGGAAGCTGCCTGCATTAATTTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
```

QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGGACTGGCGCTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGGACTGGCGCTGGAGCATC 600
QY 601 TGGTCGCATTGGGTTCACAGCAAAATCGCGCTGTAGCGGGCCCAATAAGTTCTGTCTCGG 660
Db 601 TGGTCGCATTGGGTTCACAGCAAAATCGCGCTGTAGCGGGCCCAATAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGCTGCGTGGCTGGCATTAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTGCGTGGCTGGCATTAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCACTGGAGTGCATGTCCGGTTTTCAACAACCAATGCAATGCTGA 780
Db 721 CGGAACGGGAAGGCACTGGAGTGCATGTCCGGTTTTCAACAACCAATGCAATGCTGA 780
QY 781 ATGAGGGCATCTGCCACTGCGATGCTGTGTGCCAAGATCAGATGGCGCTGGCGCGCA 840
Db 781 ATGAGGGCATCTGCCACTGCGATGCTGTGTGCCAAGATCAGATGGCGCTGGCGCGCA 840
QY 841 TGGCGGCCATTACCAAGTCCGGCTGCGGCTGGGTGGGTGCGGATATCTCGGTAGTGGATACG 900
Db 841 TGGCGGCCATTACCAAGTCCGGCTGCGGCTGGGTGGGTGCGGATATCTCGGTAGTGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACATCAAAAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACATCAAAAGGATTTTC 960
QY 961 GCCTCTGGGGCAACACGAGCTGGACCGCTTCTGCAACTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTCTGGGGCAACACGAGCTGGACCGCTTCTGCAACTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGGCCCGTCTCACTGGTGAAGAAAGAAACCCACCTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCCGTCTCACTGGTGAAGAAAGAAACCCACCTGGCGGCCAATA 1080
QY 1081 CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTAAATGCACTGGCACACAGGTTT 1140
Db 1081 CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTAAATGCACTGGCACACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGATGAGCGCAACGCAATTAATGAGTTAGTCTCACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGATGAGCGCAACGCAATTAATGAGTTAGTCTCACTCATTAG 1200
QY 1201 GCACAATCTCATGTTTGACAGCTTATCATCACTGCACGGTGCAACCAATGCTCTGGCG 1260
Db 1201 GCACAATCTCATGTTTGACAGCTTATCATCACTGCACGGTGCAACCAATGCTCTGGCG 1260
QY 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCGTGTGCAGGTTCGTAATCACTGCATTAATCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCGTGTGCAGGTTCGTAATCACTGCATTAATCG 1320
QY 1321 TGTCTGCTCAAGCGCATCTCCGTTCTGGATAATGTTTTTGGCCGACATCAATACGGTT 1380
Db 1321 TGTCTGCTCAAGCGCATCTCCGTTCTGGATAATGTTTTTGGCCGACATCAATACGGTT 1380
QY 1381 CTGGCAAAATATCTGAAATGAGCTGTGCAAAATTAATCACTCGGCTCGTAAATGTTGGA 1440
Db 1381 CTGGCAAAATATCTGAAATGAGCTGTGCAAAATTAATCACTCGGCTCGTAAATGTTGGA 1440
QY 1441 ATTGTAGCGGATAACAATTTTCAACAGAAACAGCCAGTCCGTTAGGTGTTTTCACGA 1500
Db 1441 ATTGTAGCGGATAACAATTTTCAACAGAAACAGCCAGTCCGTTAGGTGTTTTCACGA 1500
QY 1501 GCACCTCACCAAGGACCATAGATTATGAACCTGAAGAGGTAAACTGCTAATCTCG 1560
Db 1501 GCACCTCACCAAGGACCATAGATTATGAACCTGAAGAGGTAAACTGCTAATCTCG 1560
QY 1561 ATTAAACGGCGATAAAGGCTATAACCGTCTCGTGAAGTCGGTAAGAAATTCAGAGAAAGAT 1620
Db 1561 ATTAAACGGCGATAAAGGCTATAACCGTCTCGTGAAGTCGGTAAGAAATTCAGAGAAAGAT 1620
QY 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAACTCGGAAGAGAAATTCACACAGGTT 1680

Db 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGATAAACTGGAAGAGAAATTCACACAGGTT 1680
QY 1681 CGCGCAACTGCGGATGGCCCTGCACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGCGCAACTGCGGATGGCCCTGCACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTGGCTGGAATCAACCCGGACAAGAGGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTGGCTGGAATCAACCCGGACAAGAGGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGGCGTACGTTTACAACGCGAAGCTGATTGCTTACCCGATCCCTGTT 1860
Db 1801 CCGTTTACCTGGGATGGCGTACGTTTACAACGCGAAGCTGATTGCTTACCCGATCCCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCCGCCAAAACTTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCCGCCAAAACTTGGGAA 1920
QY 1921 GAGATCCCGGCTGGATAAAGAACTGAAGGAAAGTGAAGGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCTGGATAAAGAACTGAAGGAAAGTGAAGGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCCACTGGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCCACTGGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGGCAAGTACGACATTAAAGACGTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGGCAAGTACGACATTAAAGACGTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTGGTTGACATTAATAAACAACATGATGATGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTGGTTGACATTAATAAACAACATGATGATGAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTAATAAGCGCAACAGCGATGACCATCAACGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAGCGCAACAGCGATGACCATCAACGCGCCGCTGG 2220
QY 2221 GCATGGTCCAAATCGACACACAGCAAAAGTGAATTAATGGTTAAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACACAGCAAAAGTGAATTAATGGTTAAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGTTCAACCATCAAAACCGTTCTGGCGTGTGAGCGCAGGTATTAAACGCCGCCAGT 2340
Db 2281 AAGGTTCAACCATCAAAACCGTTCTGGCGTGTGAGCGCAGGTATTAAACGCCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAAGCGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
Db 2401 GAAAGCGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCACGTTATGTCGCCCAACCATGGAAGAAACGCCCAAGAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATGTCGCCCAACCATGGAAGAAACGCCCAAGAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCACGCGTCTGACTGATGATGAAGCCCTGAAAGACGCGCAGACTAATTTTCGAGCTCG 2640
Db 2581 GCACGCGTCTGACTGATGATGAAGCCCTGAAAGACGCGCAGACTAATTTTCGAGCTCG 2640
QY 2641 AACAAACAAACAAATAACAACTTCGGATTCGAGGGAAGGATTTTTCAGAAATTC 2700
Db 2641 AACAAACAAACAAATAACAACTTCGGATTCGAGGGAAGGATTTTTCAGAAATTC 2700
QY 2701 GGATCTCTTCTCTGTGGCCCGCGGCTTCGAGAGCCGG 2741

Db 2701 GGATCCTCCAGTCCCGCAAGCGCGCGCTGGTGGAGG 2741

RESULT 7
US-10-149-472-5
; Sequence 5, Application US/10149472
; Publication No. US2004002904A1
; GENERAL INFORMATION:
; APPLICANT: COUBIN-GRAMATICA, FRANCOISE
; APPLICANT: DUCOMMUN, BERNARD
; APPLICANT: PREVOST, GREGOIRE
; TITLE OF INVENTION: METHOD FOR OBTAINING HUMAN CDC25 PHOSPHATASES AND METHOD
; FILE REFERENCE: bml-427.065
; CURRENT APPLICATION NUMBER: US/10/149,472
; PRIOR APPLICATION NUMBER: PCT/FR00/03496
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: FR 99/15722
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: FR 00/06883
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: FR 00/12008
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 8101
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: JM109 / pMAL - Hs Cdc25C strain
US-10-149-472-5

Query Match 82.0%; Score 2705.8; DB 17; Length 8101;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAGAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
DB 1 CCGACACCATCGAATGGTGCAGAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCGAATTCAGGGTGGTGAATGTGAAGAACAGTAAACAGTTATACGATGTGCGAGAGTATGCCG 120
DB 61 GTCGAATTCAGGGTGGTGAATGTGAAGAACAGTAAACAGTTATACGATGTGCGAGAGTATGCCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCGAGTAACTTCCCAACCGCGTGGCAC 180
DB 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCGAGTAACTTCCCAACCGCGTGGCAC 180
QY 181 CGCGGGAAAAAGTGGAAAGCGCGGATGGCGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
DB 181 CGCGGGAAAAAGTGGAAAGCGCGGATGGCGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGAACAGTCGTGTGATGGCGTGGCACTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGGAACAGTCGTGTGATGGCGTGGCACTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTTCGCAATTTGTTCGCGCGATTAATCTCTCGCGCGATCAACTGGGTGCCAGCG 360
DB 301 ACGCGCGTTCGCAATTTGTTCGCGCGATTAATCTCTCGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTTCGATGGTGAAGACGAGCGGTGGAAGCGCTGTAAGCGGCGGTGCACAATC 420
DB 361 TGGTGGTTCGATGGTGAAGACGAGCGGTGGAAGCGCTGTAAGCGGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGGTTCAGTGGGTGATCAATTAATCTCGCTGGATGACAGATGCCA 480
DB 421 TTCTCGCGCAACCGGTTCAGTGGGTGATCAATTAATCTCGCTGGATGACAGATGCCA 480
QY 481 TTGCTGGAGGCTCCTGCACCTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
DB 481 TTGCTGGAGGCTCCTGCACCTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540

Db 481 TTGCTGGAGGCTCCTGCACCTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
QY 541 CACCCATCAACAGTATTTATTTTCCCATGAAGACGGTACGCGACTGGCGCTGGAGCATC 600
DB 541 CACCCATCAACAGTATTTATTTTCCCATGAAGACGGTACGCGACTGGCGCTGGAGCATC 600
QY 601 TGGTGCATTTGGGTACACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTCTCGG 660
DB 601 TGGTGCATTTGGGTACACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTCTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
DB 661 CGCGTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
QY 721 CGGAACCGGAAGGCGACTGGAGTCCCATGTCGGTTCCTCAACCAACCATGCAATGCTGA 780
DB 721 CGGAACCGGAAGGCGACTGGAGTCCCATGTCGGTTCCTCAACCAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTGCGATGCTGTGCGCAACGATCAGATGGCGTGGCGCAA 840
DB 781 ATGAGGGCATCGTTCCCACTGCGATGCTGTGCGCAACGATCAGATGGCGTGGCGCAA 840
QY 841 TGCAGCGCATTAACGAGTCCGGCTGCGCTGCGGATATCTCGTGGATATCTCGTGGATACG 900
DB 841 TGCAGCGCATTAACGAGTCCGGCTGCGCTGCGGATATCTCGTGGATATCTCGTGGATACG 900
QY 901 ACGATACCGAAGACAGTCAATGTTATATCCCGCTTAACCAACCATGCAATGCTGA 960
DB 901 ACGATACCGAAGACAGTCAATGTTATATCCCGCTTAACCAACCATGCAATGCTGA 960
QY 961 GCCTGCTGGGGCAAAACAGCGTGGCGCTGCGCTGCGGATATCTCGTGGATATCTCGTGGATACG 1020
DB 961 GCCTGCTGGGGCAAAACAGCGTGGCGCTGCGCTGCGGATATCTCGTGGATATCTCGTGGATACG 1020
QY 1021 AGGCAATCAGCTGTTGCGCTCTCACTGGTGAAGAAAAACACCTGGCGCCCAATA 1080
DB 1021 AGGCAATCAGCTGTTGCGCTCTCACTGGTGAAGAAAAACACCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGCAAGTGGCAACAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGCAAGTGGCAACAGGTTT 1140
QY 1141 CCGCACTGGAAGACGGCGAGTGAAGCGCAACGCAATTAATGAGTGAAGTCACTCAATAG 1200
DB 1141 CCGCACTGGAAGACGGCGAGTGAAGCGCAACGCAATTAATGAGTGAAGTCACTCAATAG 1200
QY 1201 GCACAATTCATGTTGACAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260
DB 1201 GCACAATTCATGTTGACAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGCAGCATCGGAAGCTGGTATGGCTGCGAGTGGTAAATCACTGCATTAATTCG 1320
DB 1261 TCAGGCAGCATCGGAAGCTGGTATGGCTGCGAGTGGTAAATCACTGCATTAATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGCTTCTGGATTAATGTTTTTGGCGGACATCAACCGGTT 1380
DB 1321 TGTGCTCAAGCGCACTCCCGCTTCTGGATTAATGTTTTTGGCGGACATCAACCGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGCTGGA 1440
DB 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGCTGGA 1440
QY 1441 ATTGTGAGCGGATAACAAATTTTCACAGGAACAGCGCTGCTTTAGGTGTTTTTCAACA 1500
DB 1441 ATTGTGAGCGGATAACAAATTTTCACAGGAACAGCGCTGCTTTAGGTGTTTTTCAACA 1500
QY 1501 GCATTTCAACAAAGGACCATAGATTATGAAGAACTGAAGAGGTAAATCGTGAATCTCG 1560
DB 1501 GCATTTCAACAAAGGACCATAGATTATGAAGAACTGAAGAGGTAAATCGTGAATCTCG 1560
QY 1561 ATTACGGCGATAAAGGCTATTAACCGTCTCGTGAAGTGGTAAAGAAATTCGAGAAGAT 1620
DB 1561 ATTACGGCGATAAAGGCTATTAACCGTCTCGTGAAGTGGTAAAGAAATTCGAGAAGAT 1620

QY 1621 ACCGGAATTAAAGTCAACCGTTGAGCATCCGGATAAACTGGGAAGAGAAAATTCCCAAGGTT 1680
Db 1621 ACCGGAATTAAAGTCAACCGTTGAGCATCCGGATAAACTGGGAAGAGAAAATTCCCAAGGTT 1680
QY 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACGCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACGCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTTGGCTGAATCAACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAATCAACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTAGCTTTAAGCGCAAGCTGATTGCTTACCGGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTAGCTTTAAGCGCAAGCTGATTGCTTACCGGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGTCGCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGTCGCGAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTCGATAAGAACTGAAGCGAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTCGATAAGAACTGAAGCGAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGGTGATTGCTGTGACGGGGGTTATGCTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGGTGATTGCTGTGACGGGGGTTATGCTTCAAG 2040
QY 2041 TATGAAGACGCAAGTACGACATTAAGACGTGGCGGTGATTAAGCTGGCGGAAAGG 2100
Db 2041 TATGAAGACGCAAGTACGACATTAAGACGTGGCGGTGATTAAGCTGGCGGAAAGG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAAGAAACAAACATCAATCAATCAATCAATCAAT 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAAGAAACAAACATCAATCAATCAATCAATCAAT 2160
QY 2161 TCCATCGCAAGAGTGCCTTTAATAAGCGCAAGACAGCGATCAACCGCCCGTGG 2220
Db 2161 TCCATCGCAAGAGTGCCTTTAATAAGCGCAAGACAGCGATCAACCGCCCGTGG 2220
QY 2221 GCATGGTCCACATCGACACGACGAAAGTAAATGATGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCACATCGACACGACGAAAGTAAATGATGTAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCAAAACCGTTCTGTCGCGCTGAGCGAGGTATTAAACGCCCGCAGT 2340
Db 2281 AAGGGTCAACCATCAAAACCGTTCTGTCGCGCTGAGCGAGGTATTAAACGCCCGCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGACTGATGAAGGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGTATTTCGCGCCACCATGGAAGAACCCGAGAAAGTCAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTATTTCGCGCCACCATGGAAGAACCCGAGAAAGTCAATCATG 2520
QY 2521 CCGAACAATCCCGAGATGTCGCTTTCTGGTATGCCGTGCTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACAATCCCGAGATGTCGCTTTCTGGTATGCCGTGCTGCTGCTGATCAACGCC 2580
QY 2581 GCCAGGGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2640
Db 2581 GCCAGGGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2640
QY 2641 AACCAACAACAATAAACAATAAACAATAAACAATAAACAATAAACAATAAACAATAAACA 2700
Db 2641 AACCAACAACAATAAACAATAAACAATAAACAATAAACAATAAACAATAAACAATAAACA 2700

QY 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709
RESULT 8
US-10-343-859-8
; Sequence 8, Application US/10343859
; Publication No. US20040110161A1
; GENERAL INFORMATION:
; APPLICANT: Nanogen Recognomics GMBH
; TITLE OF INVENTION: Method for detecting mutations in
; TITLE OF INVENTION: nucleotide sequences
; FILE REFERENCE: 612,406-033
; CURRENT APPLICATION NUMBER: US/10/343,859
; CURRENT FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: PCT/EP01/08127
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 10038237.1
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 6648
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: plasmid
; OTHER INFORMATION: pMAL-c2x
US-10-343-859-8

Query Match 81.8%; Score 2701; DB 19; Length 6648;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2704; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CGGACACCATCGAATGTTGCAAAACCTTTCGGGTATGGCATGATAGCCCGGAAGAGA 60
Db 1 CGGACACCATCGAATGTTGCAAAACCTTTCGGGTATGGCATGATAGCCCGGAAGAGA 60
QY 61 GTCAATTCAAGGTGTTGAATGTGAACCCAGTAACGTTATAGCATGTCGACAGATGCGG 120
Db 61 GTCAATTCAAGGTGTTGAATGTGAACCCAGTAACGTTATAGCATGTCGACAGATGCGG 120
QY 121 GTGTCCTTTATCAGACCGTTTCCGCGTGGTGAACACGAGCCAGCCAGCTTTCTGCAAAA 180
Db 121 GTGTCCTTTATCAGACCGTTTCCGCGTGGTGAACACGAGCCAGCCAGCTTTCTGCAAAA 180
QY 181 CGCGGAAAAAGTGAAGCGGCGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGGCGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC 240
QY 241 AACAACTCGCGGCAAAACAGTCGTTGTTGGGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTCGCGGCAAAACAGTCGTTGTTGGGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGGCGCTCGCAAAATGTCGCGCGGATTAATCTCGGCGCGATCAATCGGCGTGCAGG 360
Db 301 ACGGCGCTCGCAAAATGTCGCGCGGATTAATCTCGGCGCGATCAATCGGCGTGCAGG 360
QY 361 TGGTGGTTCGATGTTGAAGCAAGCGGCGTGAAGCCCTGTAAGCGGCGGTGCAATTC 420
Db 361 TGGTGGTTCGATGTTGAAGCAAGCGGCGTGAAGCCCTGTAAGCGGCGGTGCAATTC 420
QY 421 TTCTCGCGCAACGCGTCAGTGGCTGATCAATTAATCTCGGCGGTGATGATGATGATGATG 480
Db 421 TTCTCGCGCAACGCGTCAGTGGCTGATCAATTAATCTCGGCGGTGATGATGATGATGATG 480
QY 481 TTGCTGTGGAAGTGCCTGCACATTAATGTTCCGGCGTTATTTCTTGATGTTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGTGCCTGCACATTAATGTTCCGGCGTTATTTCTTGATGTTCTGACCCAGA 540
QY 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACCGGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACCGGACTGGGCGTGGAGCATC 600

Db 541 CACCCATCAACAGATAATATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTCGCAATTTGGGTCAACAGAAATCGCGCTGTGTAGCGGCGCCATTAAGCTTCTGTCTCGG 660
Db 601 TGGTCGCAATTTGGGTCAACAGAAATCGCGCTGTGTAGCGGCGCCATTAAGCTTCTGTCTCGG 660
Qy 661 CGGCTCTGGCTGTGGCTGGCTGGCTGCAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGGCTCTGGCTGTGGCTGGCTGGCTGCAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGCGCACTGAGTGGCATGTCGGTTTCCAAACCGATGCGCTGGCGCA 840
Db 721 CGGAACGGGAAGCGCACTGAGTGGCATGTCGGTTTCCAAACCGATGCGCTGGCGCA 840
Qy 781 ATGAGGGCATCTGTCCTCCATCGCATGCTGGTTGCCAAACGATCAGATGCGCTGGCGCA 840
Db 781 ATGAGGGCATCTGTCCTCCATCGCATGCTGGTTGCCAAACGATCAGATGCGCTGGCGCA 840
Qy 841 TGGCGGCCATACCGAGTCCGGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGGCCATACCGAGTCCGGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATCTGTATATCCCGCGTTAACCCACATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATCTGTATATCCCGCGTTAACCCACATCAAAACAGGATTTTC 960
Qy 961 GCCTGCTGGGCAACACGAGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGCGCGTCTCACTGTGTAAGAAAGAAACCACTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGCGCGTCTCACTGTGTAAGAAAGAAACCACTGGCGGCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGGTGGCGGATTCATTAATGACGCTGACGACGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTGGCGGATTCATTAATGACGCTGACGACGAGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGGCAACGCAATTAATGAGTTAGCTCACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGGCAACGCAATTAATGAGTTAGCTCACTCATTAG 1200
Qy 1201 GCACAATTTCTATGTTTACAGCTTATCATGACCTGACGCTGACCAATGCTTCTCGCG 1260
Db 1201 GCACAATTTCTATGTTTACAGCTTATCATGACCTGACGCTGACCAATGCTTCTCGCG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGGTATGCTGTGCGAGTCTGTAATCACTGCATAATTCTG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGCTGTGCGAGTCTGTAATCACTGCATAATTCTG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTCTCGATAATGTTTTCGCGGACATCATACCGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTCGATAATGTTTTCGCGGACATCATACCGGTT 1380
Qy 1381 CTGCAAAATTTCTGAATGAGCTGTTGACAATTAATCATCGGCTCGTAAATGTTGGA 1440
Db 1381 CTGCAAAATTTCTGAATGAGCTGTTGACAATTAATCATCGGCTCGTAAATGTTGGA 1440
Qy 1441 ATTGTAGCGGATTAACAAATTTCAACAGGAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Db 1441 ATTGTAGCGGATTAACAAATTTCAACAGGAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Qy 1501 GCATCTCACCAACAGGACCATAGATTAATGAACCTGAAGAGTAACCTGGTAATCTGG 1560
Db 1501 GCATCTCACCAACAGGACCATAGATTAATGAACCTGAAGAGTAACCTGGTAATCTGG 1560
Qy 1561 ATTAAACGGCGATAAAGGCTATAACCGTCTCGTGAAGTTCGGAAGAAATTCGAGAAGAT 1620
Db 1561 ATTAAACGGCGATAAAGGCTATAACCGTCTCGTGAAGTTCGGAAGAAATTCGAGAAGAT 1620
Qy 1621 ACCGGAATTAAGTACCGTTGAGCATCCGGATAAATCGGAAGAAATTCGAGGTT 1680
Db 1621 ACCGGAATTAAGTACCGTTGAGCATCCGGATAAATCGGAAGAAATTCGAGGTT 1680

Qy 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGACAAAGGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGACAAAGGTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACGGCAAGCTGATTTGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACGGCAAGCTGATTTGCTTACCCGATCGCTGTT 1860
Qy 1861 GAAGCGTTATGCTGATTTATAACAAAGATCTGTCGCGAACCCTCCGCAAAACCTGGGA 1920
Db 1861 GAAGCGTTATGCTGATTTATAACAAAGATCTGTCGCGAACCCTCCGCAAAACCTGGGA 1920
Qy 1921 GAGATCCCGGCGCTGGATTAAGAACTCAAAAGCGAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATTAAGAACTCAAAAGCGAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTTCACTGCGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACTGCGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAACCGCAAGTACGACATTAAGACGTTAAAGACGTTGGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACCGCAAGTACGACATTAAGACGTTAAAGACGTTGGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGTTGACCTGATTTAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTTAAACAAACACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAAGACGTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCCCGTGG 2220
Db 2161 TCCATCGCAAGACGTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCCCGTGG 2220
Qy 2221 GCATGTTCCAAACATCGACACAGCAAGTGAATTTATGTTGTTACCGTACTGCGGACCTTC 2280
Db 2221 GCATGTTCCAAACATCGACACAGCAAGTGAATTTATGTTGTTACCGTACTGCGGACCTTC 2280
Qy 2281 AAGGTCACCATCCAAAACCGTTTGGCTGCTGAGCGCAGGTATTAAACGCCCGCAGT 2340
Db 2281 AAGGTCACCATCCAAAACCGTTTGGCTGCTGAGCGCAGGTATTAAACGCCCGCAGT 2340
Qy 2341 CCGAACAAAGAGCTGCAAAAGAGTTTCTCGAAGAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGCAAAAGAGTTTCTCGAAGAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGTTTAAATAAGACAAAACCGTTGGCTGCTGAGCGCAGGTATTAAACGCCCGCAGT 2460
Db 2401 GAAGCGTTTAAATAAGACAAAACCGTTGGCTGCTGAGCGCAGGTATTAAACGCCCGCAGT 2460
Qy 2461 TTGGCGAAAGATCCAGTATTGCGCGCACCATGGAAGAAACGCCGAGAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTGCGCGCACCATGGAAGAAACGCCGAGAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTATGCGTGGCTGCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTATGCGTGGCTGCTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGATGATGCTGATGAAGCCTGAAAGAGCGCGCAGACTAAATTCGAGTCTG 2640
Db 2581 GCCAGCGTCTGATGATGCTGATGAAGCCTGAAAGAGCGCGCAGACTAAATTCGAGTCTG 2640
Qy 2641 AACAAACAAACAAATAACAAACAACTCCGGATCGAGGAGAGGATTTTCAGAAATTC 2700
Db 2641 AACAAACAAACAAATAACAAACAACTCCGGATCGAGGAGAGGATTTTCAGAAATTC 2700
Qy 2701 GGATCTCT 2709
Db 2701 GGATCTCT 2709

RESULT 9

US-10-343-859-9
; Sequence 9, Application US/10343859
; Publication No. US20040110161A1
; GENERAL INFORMATION:
; APPLICANT: Nanogen Recogonics GMBH
; TITLE OF INVENTION: Method for detecting mutations in
; FILE OF INVENTION: nucleotide sequences
; FILE REFERENCE: 612,406-033
; CURRENT APPLICATION NUMBER: US/10/343,859
; CURRENT FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: PCI/EP01/08127
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 10038237.1
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 9191
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: pMALc2x-mutS
; OTHER INFORMATION: Plasmid
US-10-343-859-9

Query Match 81.8%; Score 2698; DB 19; Length 9191;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2701; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY	1	CGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGCGCATGATAGCGCCCGGAAGAGA	60
DB	1	CGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGCGCATGATAGCGCCCGGAAGAGA	60
QY	61	GTCAATTCAGGGTGGTGAATGTGAACCACTGAACGTTATACGATGTCGAGAGTATGCGG	120
DB	61	GTCAATTCAGGGTGGTGAATGTGAACCACTGAACGTTATACGATGTCGAGAGTATGCGG	120
QY	121	GTGCTCTTATCAGACCGGTTTCCGGTGGTGAACCACTGAACGTTTCTGCGAATA	180
DB	121	GTGCTCTTATCAGACCGGTTTCCGGTGGTGAACCACTGAACGTTTCTGCGAATA	180
QY	181	CGCGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTTACATTTCCAAACCGCGTGGCAC	240
DB	181	CGCGGAAAAAGTGAAGCGCGATGCGGAGCTGAATTTACATTTCCAAACCGCGTGGCAC	240
QY	241	AACAACCTGGCGGCAACAGTCTGCTGATTTGGCGTTGCGACCTCCAGTCTGGCCCTGC	300
DB	241	AACAACCTGGCGGCAACAGTCTGCTGATTTGGCGTTGCGACCTCCAGTCTGGCCCTGC	300
QY	301	ACGCGCGTTCGCAAAATTTGTCGGCGGATTAATAATCTCGCGCGATCAACTGGGTGCCAGG	360
DB	301	ACGCGCGTTCGCAAAATTTGTCGGCGGATTAATAATCTCGCGCGATCAACTGGGTGCCAGG	360
QY	361	TGGTGGTTCGATGGTAGAACGAGCGCGTTCGAAGCCTGTAAAGCGGCGTGCACAAATC	420
DB	361	TGGTGGTTCGATGGTAGAACGAGCGCGTTCGAAGCCTGTAAAGCGGCGTGCACAAATC	420
QY	421	TTCTCGGCAACCGCTCAGTGGGTGATCATTAATCTATCCGCTGGATGACAGGATGCCA	480
DB	421	TTCTCGGCAACCGCTCAGTGGGTGATCATTAATCTATCCGCTGGATGACAGGATGCCA	480
QY	481	TTGCTGTGAAGTGCCTCACTAATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA	540
DB	481	TTGCTGTGAAGTGCCTCACTAATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA	540
QY	541	CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGGACTTGGCGGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGGACTTGGCGGTGGAGCATC	600
QY	601	TGGTGGATGGGTGCACAGCAAAATCGCGCTGTAGCGGGCCATTAAGTTCTGTCTCGG	660
DB			

DB	601	TGGTGGATGGGTGCACAGCAAAATCGCGCTGTGTAGCGGGCCCATTAAGTTCTGTCTCGG	660
QY	661	CGCGTCTCGCTCGCTCGGTCATGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	661	CGCGTCTCGCTCGCTCGGTCATGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAACGGGAAGGCGACTGGAAGTGCATGTCGGTTTCAACAAACCATGCAATTCGTGA	780
DB	721	CGGAACGGGAAGGCGACTGGAAGTGCATGTCGGTTTCAACAAACCATGCAATTCGTGA	780
QY	781	ATGAGGCGATCTGTTCCCACTCGCATGCTGTTGCCAAGATCAGATGCGCTGGCGGCA	840
DB	781	ATGAGGCGATCTGTTCCCACTCGCATGCTGTTGCCAAGATCAGATGCGCTGGCGGCA	840
QY	841	TGCGGCCATTACCGAGTCCGGGCTCGCGTGGTGGGATATCTCGGTAGTGGGATAGC	900
DB	841	TGCGGCCATTACCGAGTCCGGGCTCGCGTGGTGGGATATCTCGGTAGTGGGATAGC	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAACATCAACAGGATTTTC	960
DB	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAACATCAACAGGATTTTC	960
QY	961	GCCTCTGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCAGGCGGTGA	1020
DB	961	GCCTCTGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCAGGCGGTGA	1020
QY	1021	AGGCAATCAGCTGTTGCCCGCTCTCACTGCTGAAGAAAAACCAACCTGGCGCCCAATA	1080
DB	1021	AGGCAATCAGCTGTTGCCCGCTCTCACTGCTGAAGAAAAACCAACCTGGCGCCCAATA	1080
QY	1081	CGAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGCTGGCAGCAGAGTTT	1140
DB	1081	CGAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGCTGGCAGCAGAGTTT	1140
QY	1141	CCGACTCGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTAGTCTACTCATTTAG	1200
DB	1141	CCGACTCGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTAGTCTACTCATTTAG	1200
QY	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCGGTCGACCAATGCTCTGGCG	1260
DB	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCGGTCGACCAATGCTCTGGCG	1260
QY	1261	TCAGGCGCATCGAAGCTGTGTATGCTGTGCGAGTGTGCAATCACTGATTAATTCG	1320
DB	1261	TCAGGCGCATCGAAGCTGTGTATGCTGTGCGAGTGTGCAATCACTGATTAATTCG	1320
QY	1321	TGTCGCTCAAGCGCACCTCCGTTCTGATTAATGTTTTCGCGCGACATCAACGGTT	1380
DB	1321	TGTCGCTCAAGCGCACCTCCGTTCTGATTAATGTTTTCGCGCGACATCAACGGTT	1380
QY	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGGA	1440
DB	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGGA	1440
QY	1441	ATTGTGAGCGGATAACAATTTTCAACAGAAACAGCCAGTCCGTTAGGTGTTTTCACGA	1500
DB	1441	ATTGTGAGCGGATAACAATTTTCAACAGAAACAGCCAGTCCGTTAGGTGTTTTCACGA	1500
QY	1501	GCACCTTCAACAGGACCATAGATTATGAACTGGAAGAGGTAACTGCTGTAATCTCG	1560
DB	1501	GCACCTTCAACAGGACCATAGATTATGAACTGGAAGAGGTAACTGCTGTAATCTCG	1560
QY	1561	ATTAAACGGCGATAAAGGCTATAACCGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
DB	1561	ATTAAACGGCGATAAAGGCTATAACCGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAATTCGGAAGAGAAATTCACAGGTT	1680
DB	1621	ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAATTCGGAAGAGAAATTCACAGGTT	1680
QY	1681	GGGCAACTGGCGATGGCCCTGACATTATCTCTGGGCACACGCGCTTCTGGCTAC	1740
DB	1681	GGGCAACTGGCGATGGCCCTGACATTATCTCTGGGCACACGCGCTTCTGGCTAC	1740


```
QY 1741 GCTCAATCTGGCTGTGGCTGAAATACCCCGGACAAAGCGTTCAGGACAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATACCCCGGACAAAGCGTTCAGGACAGCTGTAT 1800
QY 1801 CGTTTACCTGGATGCGGTAGCTTAAACCGCAAGCTGATGCTTACCGGATCGTGT 1860
Db 1801 CGTTTACCTGGATGCGGTAGCTTAAACCGCAAGCTGATGCTTACCGGATCGTGT 1860
QY 1861 GAAGCGTATCGTGTATTAACCAAGATCTGCTGCGCAACCGCCCAAAACCTTGGGAA 1920
Db 1861 GAAGCGTATCGTGTATTAACCAAGATCTGCTGCGCAACCGCCCAAAACCTTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATGAAGAACTGAAGCGCAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATGAAGAACTGAAGCGCAAGGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAGACGGCAAGTACGACATTAAGACGTTGGCGGTGATTAACCGCGGAAACCG 2100
Db 2041 TATGAAGACGGCAAGTACGACATTAAGACGTTGGCGGTGATTAACCGCGGAAACCG 2100
QY 2101 GGTCTGACCTTCTGCTGATGCTGATTAAGAAACCAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGATGCTGATTAAGAAACCAACACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTCCCTTAAATAAGCGGAAACAGCGATGACCATCAACGGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTCCCTTAAATAAGCGGAAACAGCGATGACCATCAACGGCCGCTGG 2220
QY 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAAGAAACCAACACATGAATGCAGACACCGATTAC 2280
Db 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAAGAAACCAACACATGAATGCAGACACCGATTAC 2280
QY 2281 AAGGTCAACCATCAACACCGTTCGTCGCTGCTGAGCGAGGTATTAACCGCGCCAGT 2340
Db 2281 AAGGTCAACCATCAACACCGTTCGTCGCTGCTGAGCGAGGTATTAACCGCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGGTTCTCGAAACTATCTGCTGATGAAGGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGGTTCTCGAAACTATCTGCTGATGAAGGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGTGGGTGCTGAGCGAGGTATTAACCGCGCCAGT 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGTGGGTGCTGAGCGAGGTATTAACCGCGCCAGT 2460
QY 2461 TTGGGGAAGATCCACGTATTTGCGCCACCAATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGGGAAGATCCACGTATTTGCGCCACCAATGGAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTATCGCGTGGTATCGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTATCGCGTGGTATCGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACAGTCTGATGAAGCCCTGAAAGACCGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTCTGATGAAGCCCTGAAAGACCGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACCAACACACATTAACATTAACACACCTCGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACCAACACACATTAACATTAACACACCTCGGATCGAGGGAAGGATTCAGAAATTC 2700
QY 2701 GGATCC 2706
Db 2701 GGATCC 2706
```

RESULT 10
US-10-263-153-40
./ Sequence 40, Application US/10263153

```
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Gineburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263.153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 7112
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3189)
; OTHER INFORMATION: pMBP-c2X-Toxop30del11 (52-214aa)
US-10-263-153-40
```

```
Query Match      81.5%   Score 2690.4;   DB 18;   Length 7112;
Best Local Similarity 99.8%;   Pred. No. 0;
Matches 2694;   Conservative 0;   Mismatches 6;   Indels 0;   Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTCAGGTTGGTGAATGTGAACCAAGTATACGATGTGCGAGATATGCGG 120
Db 61 GTCAATTCAGGTTGGTGAATGTGAACCAAGTATACGATGTGCGAGATATGCGG 120
QY 121 GTGCTCTTATCAGACCGTTTCCGCGGTGTGAACAGGCGGACGAGTTCCTGCGGAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCGCGGTGTGAACAGGCGGACGAGTTCCTGCGGAAA 180
QY 181 CGCGGAAAAAAGTGAAGCGGCGATGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAAGTGAAGCGGCGATGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
QY 241 AACATCGCGCGGCAACAGTCTGCTGATGCGGCTTGGCACCTCAGTCTGGCCCTGCG 300
Db 241 AACATCGCGCGGCAACAGTCTGCTGATGCGGCTTGGCACCTCAGTCTGGCCCTGCG 300
QY 301 ACGCGCGCTCGCAAAATGTCGCGCGGATTAATTCCTCGCGCGGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGCTCGCAAAATGTCGCGCGGATTAATTCCTCGCGCGGATCAACTGGGTGCCAGCG 360
QY 361 TGTGTGTGTCGATGTTAGAACGAGCGGCTGGAAGCTGTAAAGCGGCGGTGCACAATC 420
Db 361 TGTGTGTGTCGATGTTAGAACGAGCGGCTGGAAGCTGTAAAGCGGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGATGACGAGATGCGCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGATGACGAGATGCGCA 480
QY 481 TTGCTGTGAAGCTGCTGCACTTAATGTTCCGCGGCTTAATTTCTTGATGTCTCTGACGAGA 540
Db 481 TTGCTGTGAAGCTGCTGCACTTAATGTTCCGCGGCTTAATTTCTTGATGTCTCTGACGAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGTTACGCGAATGCGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGTTACGCGAATGCGGCGTGGAGCATC 600
QY 601 TGTGCGCATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGCGCCATTTAAGTCTGCTCGG 660
Db 601 TGTGCGCATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGCGCCATTTAAGTCTGCTCGG 660
QY 661 CGCGTCTGCGTCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
```

Db 661 |||||CGGCTCTGGCTGGCTGGCAATAATCTCACTCGCANATAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGACCTGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 |||||CGGAACGGGAAGGACCTGGAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGGATCTTCCCACTGGGATGCTGGTTGCCAAGATCAGATGCGCTGGCGGCA 840
Db 781 ATGAGGGATCTTCCCACTGGGATGCTGGTTGCCAAGATCAGATGCGCTGGCGGCA 840
Qy 841 TGCGGCCATTACAGATCCGGCTCGCGTGGTGGCGGATATCTCGGTAGTGGGATACG 900
Db 841 TGCGGCCATTACAGATCCGGCTCGCGTGGTGGCGGATATCTCGGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACCAACCATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACCAACCATCAAAACAGGATTTTC 960
Qy 961 GCCTCTGGGGCAACCAAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTCTGGGGCAACCAAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAACCAACCTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAGCACAGGTTT 1140
Db 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAGCACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Qy 1201 GCACAATTCATGTTTGCAGCTTATCATGACTGCAGGTGCACCAATGCTTCTGGG 1260
Db 1201 GCACAATTCATGTTTGCAGCTTATCATGACTGCAGGTGCACCAATGCTTCTGGG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGTAATCACTGCATAAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGTAATCACTGCATAAATTCG 1320
Qy 1321 TGTCTCTCAAGCGCACTCCCGTTCTGGATTAATGTTTTTGGCGCAGCATCAACGGTT 1380
Db 1321 TGTCTCTCAAGCGCACTCCCGTTCTGGATTAATGTTTTTGGCGCAGCATCAACGGTT 1380
Qy 1381 CTGGCAATATCTGAAATGAGCTGTGACATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAATATCTGAAATGAGCTGTGACATTAATCATCGGCTCGTATATGTTGGA 1440
Qy 1441 ATTGTAGCGGATAACAATTTCCACAGGAAACAGCCAGTCCGTTTAGTGTGTTTTCAG 1500
Db 1441 ATTGTAGCGGATAACAATTTCCACAGGAAACAGCCAGTCCGTTTAGTGTGTTTTCAG 1500
Qy 1501 GCATTTCAACCAAGGACCATAGATATGAAAACCGAAGAGTTAACTGGTAATCTGG 1560
Db 1501 GCATTTCAACCAAGGACCATAGATATGAAAACCGAAGAGTTAACTGGTAATCTGG 1560
Qy 1561 ATTAACGGGATTAAGGCTATTAACGGTCTCGCTGAAGTCCGTTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTAACGGGATTAAGGCTATTAACGGTCTCGCTGAAGTCCGTTAAGAAATTCAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACGTTGAGATCCGATTAACCTGGAAGAGAAATTCACACAGTT 1680
Db 1621 ACCGGAATTAAGTCAACGTTGAGATCCGATTAACCTGGAAGAGAAATTCACACAGTT 1680
Qy 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCAACAGCCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCAACAGCCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGCTGAATATCAACCGGCAACCGGTTCCAGGACAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAATATCAACCGGCAACCGGTTCCAGGACAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTTAAACCGCAAGCTGATTGTTTACCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTAAACCGCAAGCTGATTGTTTACCGATCGCTGT 1860
Qy 1861 GAAGGTTATCGCTGATTATTAACAAAGATCTGTCGCAACCGCCCAAAACCTGGGAA 1920
Db 1861 GAAGGTTATCGCTGATTATTAACAAAGATCTGTCGCAACCGCCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCTGGATAAAGAACTGAAACGAAAGGTAAAGAGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCTGGATAAAGAACTGAAACGAAAGGTAAAGAGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCTGATCTTCACTGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCTGATCTTCACTGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTAAGACGTGGGCTGGATTAACGCTGGCGGAAAGG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGTGGGCTGGATTAACGCTGGCGGAAAGG 2100
Qy 2101 GGTCTGACCTTCTGTTGACCTGATTAAAAAACAACACATGAAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTGACCTGATTAAAAAACAACACATGAAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGTTCAACATCGACACCAAGTAAATTAATGTTGTTAAACGTTACTGCGGACCTTC 2280
Db 2221 GCATGTTCAACATCGACACCAAGTAAATTAATGTTGTTAAACGTTACTGCGGACCTTC 2280
Qy 2281 AAGGTTCAACCATCAACACCGTTTGGTGGCTGTGAGCGAGGTATTAACGCCCCAGT 2340
Db 2281 AAGGTTCAACCATCAACACCGTTTGGTGGCTGTGAGCGAGGTATTAACGCCCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACCTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACCTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGTTTAAAGAACAAACCGTGGCTGGCTGAGCTGAACTCTTACGAGGAAG 2460
Db 2401 GAAGCGTTTAAAGAACAAACCGTGGCTGGCTGAGCTGAACTCTTACGAGGAAG 2460
Qy 2461 TTGGCAAGATCCACGATTTGCGCCCATATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCAAGATCCACGATTTGCGCCCATATGGAACCGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGCTGCTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGCTGCTGCTGCTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGATGATGAGCCCTGAAAGACGGCGAGACTAAATTCGAGTCTG 2640
Db 2581 GCCAGCGTCTGATGATGAGCCCTGAAAGACGGCGAGACTAAATTCGAGTCTG 2640
Qy 2641 AACCAACCAACATAACATAACCAACCTCGGGATCGAGGGAAGGATTTTCAAAATTC 2700
Db 2641 AACCAACCAACATAACATAACCAACCTCGGGATCGAGGGAAGGATTTTCAAAATTC 2700

RESULT 11

US-10-263-153-30
; Sequence 30, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.

; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
 ; FILE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
 ; FILE REFERENCE: 6984.US.O1
 ; CURRENT APPLICATION NUMBER: US/10/263,153
 ; CURRENT FILING DATE: 2002-10-22
 ; NUMBER OF SEQ ID NOS: 74
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 30
 ; LENGTH: 7259
 ; TYPE: DNA
 ; ORGANISM: Toxoplasma gondii
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1528)...(3336)
 ; OTHER INFORMATION: pMBP-C2X-Toxop30del14del18 (83-294aa)
 US-10-263-153-30

Query Match 81.5%; Score 2690.4; DB 18; Length 7259;
 Best Local Similarity 99.8%; Pred. No. 0;

Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY	1	CGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA	60
DB	1	CGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA	60
QY	61	GTCAATTCAGGGTGTGAATGTGAACCAAGTAAAGTTATACGATGTGCGAGATATGCCG	120
DB	61	GTCAATTCAGGGTGTGAATGTGAACCAAGTAAAGTTATACGATGTGCGAGATATGCCG	120
QY	121	GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACCGAGCCAGCCAGCTTTCTCGGAAA	180
DB	121	GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACCGAGCCAGCCAGCTTTCTCGGAAA	180
QY	181	CGCGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTAATTCACCAACCGCGTGGCAC	240
DB	181	CGCGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTAATTCACCAACCGCGTGGCAC	240
QY	241	AACAACTGGCGGCGMAACAGTCTGTTGCTGATGGCGTTGCCACTCCAGTCTGCCCTGC	300
DB	241	AACAACTGGCGGCGMAACAGTCTGTTGCTGATGGCGTTGCCACTCCAGTCTGCCCTGC	300
QY	301	ACGCGCGTCCGCAATTTGTCGCGCGATTAATCTCCGCGCGATCAACTGGTGCCAGCG	360
DB	301	ACGCGCGTCCGCAATTTGTCGCGCGATTAATCTCCGCGCGATCAACTGGTGCCAGCG	360
QY	361	TGGTGGTTCGATGGTAGAACGAAGCGCGTGAAGCCTGTAAAGCGCGGTGCACAATC	420
DB	361	TGGTGGTTCGATGGTAGAACGAAGCGCGTGAAGCCTGTAAAGCGCGGTGCACAATC	420
QY	421	TTCTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACGAGATGCCA	480
DB	421	TTCTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACGAGATGCCA	480
QY	481	TTGCTGTGGAGCTCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTTCTTGACCAGA	540
DB	481	TTGCTGTGGAGCTCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTTCTTGACCAGA	540
QY	541	CACCCATCAACAGTATTTATTTCTCCCATGAAGACGTAACGAGTGGCGGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTTATTTCTCCCATGAAGACGTAACGAGTGGCGGTGGAGCATC	600
QY	601	TGGTGGCATTTGGGTACACAGCAATCGCGTGTATAGCGGCGCCATTAAAGTCTCTCGG	660
DB	601	TGGTGGCATTTGGGTACACAGCAATCGCGTGTATAGCGGCGCCATTAAAGTCTCTCGG	660
QY	661	CGGCTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	661	CGGCTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAACGGGAAGCGACTGAGTGCATGTCGGTTTTTCAACAAAACCATGCAAAATGCTGA	780
DB	721	CGGAACGGGAAGCGACTGAGTGCATGTCGGTTTTTCAACAAAACCATGCAAAATGCTGA	780

QY	781	ATGAGGCATCGTTCCCACTGCGATGCTGTTGCCAAGATCAGATGGCGCTGGCGCGAA	840
DB	781	ATGAGGCATCGTTCCCACTGCGATGCTGTTGCCAAGATCAGATGGCGCTGGCGCGAA	840
QY	841	TGCGCGCATATTACCGAGTCCGGGCTCGCGTGTGGTGGGATATCTCGGTAGTGGATACG	900
DB	841	TGCGCGCATATTACCGAGTCCGGGCTCGCGTGTGGTGGGATATCTCGGTAGTGGATACG	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACATCAAAACAGATTTTC	960
DB	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACATCAAAACAGATTTTC	960
QY	961	GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020
DB	961	GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020
QY	1021	AGGCAATACAGTGTGTCGGCTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA	1080
DB	1021	AGGCAATACAGTGTGTCGGCTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA	1080
QY	1081	CGCAACCGCTCTCCCGCGGTTGGCGATTCATTAATGCGAGCTGGCAACAGAGTTT	1140
DB	1081	CGCAACCGCTCTCCCGCGGTTGGCGATTCATTAATGCGAGCTGGCAACAGAGTTT	1140
QY	1141	CCGACCTGGAAGCGGCGAGTGAACGCAATTAATGAGTTAGCTTACCTCATTTAG	1200
DB	1141	CCGACCTGGAAGCGGCGAGTGAACGCAATTAATGAGTTAGCTTACCTCATTTAG	1200
QY	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGGTGCACCAATGCTTGGCG	1260
DB	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGGTGCACCAATGCTTGGCG	1260
QY	1261	TCAGGAGCCCATCGGAGCTGTTGCTGAGTGGCTGCGAGTCTGTAATCACTGCAATTCG	1320
DB	1261	TCAGGAGCCCATCGGAGCTGTTGCTGAGTGGCTGCGAGTCTGTAATCACTGCAATTCG	1320
QY	1321	TGTCGCTCAAGCGCGCACTCCCGTCTTGGATAATGTTTTTGGCCGACATCATACCGGTT	1380
DB	1321	TGTCGCTCAAGCGCGCACTCCCGTCTTGGATAATGTTTTTGGCCGACATCATACCGGTT	1380
QY	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGCTCGTATATGTTGGA	1440
DB	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGCTCGTATATGTTGGA	1440
QY	1441	ATTGTGAGCGGATAACAAATTTTACACAGGAACAGCCAGTCCGTTTAGTGTTTTCACGA	1500
DB	1441	ATTGTGAGCGGATAACAAATTTTACACAGGAACAGCCAGTCCGTTTAGTGTTTTCACGA	1500
QY	1501	GCATTTCAACCAACAGGACCATAGATTTATGAAACTGAAAGAGGTAAACTGGTAACTCG	1560
DB	1501	GCATTTCAACCAACAGGACCATAGATTTATGAAACTGAAAGAGGTAAACTGGTAACTCG	1560
QY	1561	ATTACCGGATAAAGGCTTAAACCGTCTCGCTGAGTCCGTTAGGAAATTCGAAAGAT	1620
DB	1561	ATTACCGGATAAAGGCTTAAACCGTCTCGCTGAGTCCGTTAGGAAATTCGAAAGAT	1620
QY	1621	ACCGAATTTAAAGTCAACCGTTCAGCATCCGGATTAACCTGAAAGAGAAATTCACACAGGTT	1680
DB	1621	ACCGAATTTAAAGTCAACCGTTCAGCATCCGGATTAACCTGAAAGAGAAATTCACACAGGTT	1680
QY	1681	GGGCAACTGGGATGGCCCTGACATTTCTTCTGGGCAACAGCCGTTTGGTGCTAC	1740
DB	1681	GGGCAACTGGGATGGCCCTGACATTTCTTCTGGGCAACAGCCGTTTGGTGCTAC	1740
QY	1741	GCTCAATCTGGCTGTGGCTGAATCAACCGGAAAGCTGATGCTTACCCGATCGCTGTT	1800
DB	1741	GCTCAATCTGGCTGTGGCTGAATCAACCGGAAAGCTGATGCTTACCCGATCGCTGTT	1800
QY	1801	CGGTTTACCTGGGATGCCGATCGTTTACAAACGGCAAGCTGATGCTTACCCGATCGCTGTT	1860
DB	1801	CGGTTTACCTGGGATGCCGATCGTTTACAAACGGCAAGCTGATGCTTACCCGATCGCTGTT	1860
QY	1861	GAAAGCTTACCTGATTTATTAACAAAGATCTGCTGCCGAAACCCGCCAAAAAACCTGGGAA	1920

Db 1861 GAAGCGTTATCGCTGATTTATTAACAAGATCTGTCGCGAAACCCGCCAAAAACCTGGGAA 1920
Qy 1921 GAGATCCGCGCGCTGGATAAAGAACTGAAGAGCGAAGGTGAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGCGCGCTGGATAAAGAACTGAAGAGCGAAGGTGAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTCGCGCGCTGATGTTGCTGTCAGCGGGGGTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTCGCGCGCTGATGTTGCTGTCAGCGGGGGTATGCGTTCAAG 2040
Qy 2041 TATGAAAAACGCAAGTACACATTAAGAAGCGTGGCGTGGATAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACACATTAAGAAGCGTGGCGTGGATAACGCTGGCGCGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGTTGACCTGATTAATAAACAACATGAATGACAGACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTAATAAACAACATGAATGACAGACCGGATTAC 2160
Qy 2161 TCCATCGCAGAAGCTGCTTTAATAAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAAGCTGCTTTAATAAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
Qy 2221 GCATGGTCCAACTCGACACAGCAAGTGAATTAATGTTAAACGTTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAACTCGACACAGCAAGTGAATTAATGTTAAACGTTACTGCGGACCTTC 2280
Qy 2281 AAGGTCACCACTCAACACCGTTTCGTCGCGTGTGAGCGAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTCACCACTCAACACCGTTTCGTCGCGTGTGAGCGAGGTATTAACGCCGCCAGT 2340
Qy 2341 CCGAAACAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGATGATGAAGTCTG 2400
Db 2341 CCGAAACAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGATGATGAAGTCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGTGGTGGCGTGGAGCGTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGTGGTGGCGTGGAGCGTGAAGTCTTACGAGGAAG 2460
Qy 2461 TTGGCGAAAGATCCACGATTTCCGCCCAACCTGGAAGAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTTCCGCCCAACCTGGAAGAACGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTGCTGATGATGAAGTCAAGGCC 2580
Db 2521 CCGAATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTGCTGATGATGAAGTCAAGGCC 2580
Qy 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGGAAGCGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGGAAGCGCGCAGACTAAATTCGAGCTCG 2640
Qy 2641 AACAAACAACAATAACAATAACAACCACTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACCACTCGGGATCGAGGGAAGGATTCAGAAATTC 2700

RESULT 12
US-10-263-153-35
; Sequence 35, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Gineburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.O1
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 35
; LENGTH: 7322
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528) ... (3399)
; OTHER INFORMATION: pMBP-c2X-Toxop30del10 (52-284aa)
US-10-263-153-35

Query Match 81.5%; Score 2690.4; DB 18; Length 7322;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGGCATGTAGGCCCGGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGGCATGTAGGCCCGGGAAGAGA 60
Qy 61 GTCATTCAGGGTGGTGAATGTAACCAAGTAAAGTTATACGATGTCGAGAGTATGCCG 120
Db 61 GTCATTCAGGGTGGTGAATGTAACCAAGTAAAGTTATACGATGTCGAGAGTATGCCG 120
Qy 121 GTGTCCTTATCAGACCGCTTTCCCGCGTGTGAACCAAGCCAGCCACGTTTCTGCGAAAA 180
Db 121 GTGTCCTTATCAGACCGCTTTCCCGCGTGTGAACCAAGCCAGCCACGTTTCTGCGAAAA 180
Qy 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTACATTCCTCCAAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTACATTCCTCCAAACCGCGTGGCAC 240
Qy 241 AACAACTGGCGGGCAACAGTCGTTGATGTTGGGTTGCCACCTCCAGTCTGGGCCCTGC 300
Db 241 AACAACTGGCGGGCAACAGTCGTTGATGTTGGGTTGCCACCTCCAGTCTGGGCCCTGC 300
Qy 301 ACGCGCGTTCGCAAAATTCGCGCGGATTAATCTCGCGCGCATCAATCGGCGTGCAGCG 360
Db 301 ACGCGCGTTCGCAAAATTCGCGCGGATTAATCTCGCGCGCATCAATCGGCGTGCAGCG 360
Qy 361 TGGTGGTTCGATGATAGAACGAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCAACAATC 420
Db 361 TGGTGGTTCGATGATAGAACGAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCAACAATC 420
Qy 421 TTCTCGCGCAACGGCTAGTGGGTGATCAATTAATCTCCGTGATGATGACCAAGATGCCA 480
Db 421 TTCTCGCGCAACGGCTAGTGGGTGATCAATTAATCTCCGTGATGATGACCAAGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCGCTGCACTAATGTTCCGGGGTATTTCTTGATGTTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCGCTGCACTAATGTTCCGGGGTATTTCTTGATGTTCTGACCCAGA 540
Qy 541 CACCATCAACAGATTAATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCATCAACAGATTAATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Qy 601 TGGTGGCATTTGGGTACCGAATAATCGCGCTGTTAGCGGGGCCATTAAGATTTCTGTCGG 660
Db 601 TGGTGGCATTTGGGTACCGAATAATCGCGCTGTTAGCGGGGCCATTAAGATTTCTGTCGG 660
Qy 661 CGCGTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGAAACGGGAAGGGACCTGGAGTGCATGTCCGGTTTTTCAACAAACCAATGCAAAATGCTGA 780
Db 721 CGAAACGGGAAGGGACCTGGAGTGCATGTCCGGTTTTTCAACAAACCAATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATCGTTCCCACTCGCATGCTGGTTGTCGCAACGATCAGATGCGGTGGCGCAA 840
Db 781 ATGAGGGCATCGTTCCCACTCGCATGCTGGTTGTCGCAACGATCAGATGCGGTGGCGCAA 840
Qy 841 TGGCGGCCATTACCGAGTCCGGGCTGGCGGTTGGTGGGATATCTCGGTAGTGGGATAGC 900
Db 841 TGGCGGCCATTACCGAGTCCGGGCTGGCGGTTGGTGGGATATCTCGGTAGTGGGATAGC 900

```
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTACCCATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTACCCATCAAAACAGGATTTTC 960
QY 961 GCCTGTCTGGGGGAAACACAGGTGGACCGCTTGTGCAACTCTCTCAGGGCCACAGGGGTGA 1020
Db 961 GCCTGTCTGGGGGAAACACAGGTGGACCGCTTGTGCAACTCTCTCAGGGCCACAGGGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGTTGCGCGCTCTACCTGTGAAAGAAACACACCTTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGTTGCGCGCTCTACCTGTGAAAGAAACACACCTTGGCGGCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGAATTAATGAGCTGGCAGCACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGAATTAATGAGCTGGCAGCACAGGTTT 1140
QY 1141 CCGGACTGAAAGCGGCGAGTAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Db 1141 CCGGACTGAAAGCGGCGAGTAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAATTCATGTTGACAGCTTATCATCGACTGCAGGTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTCATGTTGACAGCTTATCATCGACTGCAGGTGACCAATGCTTCTGGCG 1260
QY 1261 TCAGGAGCGCATCGAAGCTGTGGTATGGCTGTGCAGGTTCGTAATCACTGCATAATTCG 1320
Db 1261 TCAGGAGCGCATCGAAGCTGTGGTATGGCTGTGCAGGTTCGTAATCACTGCATAATTCG 1320
QY 1321 TGTGCTCTAAGCGCACCTCCCGTTCGGTATGTTTTCGGCGGAGCATCAACCGTT 1380
Db 1321 TGTGCTCTAAGCGCACCTCCCGTTCGGTATGTTTTCGGCGGAGCATCAACCGTT 1380
QY 1381 CTGGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA 1440
Db 1381 CTGGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA 1440
QY 1441 ATTGTAGCGGATAACAATTTTACACAGGAAACAGCCAGTCCGTTTGGTGTTCACGA 1500
Db 1441 ATTGTAGCGGATAACAATTTTACACAGGAAACAGCCAGTCCGTTTGGTGTTCACGA 1500
QY 1501 GCACCTTACCAACAGGACCATAGATTATGAAACTGAAAGAGTAAACTGGTAAATCTGG 1560
Db 1501 GCACCTTACCAACAGGACCATAGCATATGAAATCGAAGAGTAAACTGGTAAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCACCGTTGACATCCGGATAACTGGNAGAGAAATTCACACAGGTT 1680
Db 1621 ACCGGAATTAAGTCACCGTTGACATCCGGATAACTGGNAGAGAGAAATTCACACAGGTT 1680
QY 1681 CGCGCAATCGCGATGGCCCTGACATATCTTCTGGGCACACGACCCCTTTGGTGGCTAC 1740
Db 1681 CGCGCAATCGCGATGGCCCTGACATATCTTCTGGGCACACGACCCCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCCTTTGGCTGAAATCACCCCGACAAAGGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCCTTTGGCTGAAATCACCCCGACAAAGGTTCCAGGACAAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGTACGTTACACCGCAAGCTGATTCGTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGTACGTTACACCGCAAGCTGATTCGTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAAACCCCGCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAAACCCCGCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
```

```
QY 1981 CTGCAAGAACCGTACTTCACTGCGCCGCTGATTCCTGACGGGGGTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCCGCTGATTCCTGACGGGGGTATGCGTTCAAG 2040
QY 2041 TATGAAACCGCAAGTAGTACGACATTAAGACGTTGGCGTGGATACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTAGTACGACATTAAGACGTTGGCGTGGATACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCTGTTGACCTGATTTAAACAAACACATGAATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTTAAACAAACACATGAATGACAGACACCGATTAC 2160
QY 2161 TCCATCCGAGNAGCTGCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCCGAGNAGCTGCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCAAACATCGACACAGCAAAAGTGAATTTATGTTGTTAAACGCTGACCTTC 2280
Db 2221 GCATGGTCCAAACATCGACACAGCAAAAGTGAATTTATGTTGTTAAACGCTGACCTTC 2280
QY 2281 AAGGTTCAACCATCCAAACCGTTTCTGTTGGCGTCTGAGCGCAGGTATTAACGGCCAGT 2340
Db 2281 AAGGTTCAACCATCCAAACCGTTTCTGTTGGCGTCTGAGCGCAGGTATTAACGGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTTCTCGAATACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTTCTCGAATACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGCGGAAAGATCCAGTATTGCGCGCACCATGGAAGAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGCGGAAAGATCCAGTATTGCGCGCACCATGGAAGAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCGTTTCTGATGCGGTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCCGCAGATGTCGCGTTTCTGATGCGGTGCTGCTGCTGCTGCTGCTGCTG 2580
QY 2581 GCAGCGGCTGCTGAGACTGTGATGAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCAGCGGCTGCTGAGACTGTGATGAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACCAACAAACAAATTAACATTAACAAACCTCGGATCGAGGAGGATTTTCAGAAATTC 2700
Db 2641 AACCAACAAACAAATTAACATTAACAAACCTCGGATCGAGGAGGATTTTCAGAAATTC 2700
```

RESULT 13
US-10-263-153-25
; Sequence 25, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Gineburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.O1
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 7352
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3429)

; OTHER INFORMATION: pMBP-c2X-ToxoP30del14C (52-294aa)									
US-10-263-153-25									
Query Match 81.5%; Score 2690.4; DB 18; Length 7352;									
Best Local Similarity 99.8%; Pred. No. 0;									
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;									
QY	1	CGGACCAATCAATGGTGGCAAAACCTTTCCGGGTATGGCATGATAGGCCCGCGGAAGAGA	60						
DB	1	CGGACCAATCAATGGTGGCAAAACCTTTCCGGGTATGGCATGATAGGCCCGCGGAAGAGA	60						
QY	61	GTCAATTCAGGGTGGTGAATGTGAACAGTAACTATACGATCTCCAGAGTATCGG	120						
DB	61	GTCAATTCAGGGTGGTGAATGTGAACAGTAACTATACGATCTCCAGAGTATCGG	120						
QY	121	GTGTCCTTATCAGACCCGTTTCCCGGTGGTGAACAGGCCAGCCACGCTTCTCGGAAAA	180						
DB	121	GTGTCCTTATCAGACCCGTTTCCCGGTGGTGAACAGGCCAGCCACGCTTCTCGGAAAA	180						
QY	181	CGCGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTTACATTCCTCAACCCGCTGGCAC	240						
DB	181	CGCGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTTACATTCCTCAACCCGCTGGCAC	240						
QY	241	AACAACTGCGGGCAAAACAGTCTGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300						
DB	241	AACAACTGCGGGCAAAACAGTCTGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300						
QY	301	ACGCGCCGTCGCAAAATTTGCGCGCGATTAAATCTCGCGCGGATCAACTGGGTGCCAGCG	360						
DB	301	ACGCGCCGTCGCAAAATTTGCGCGCGATTAAATCTCGCGCGGATCAACTGGGTGCCAGCG	360						
QY	361	TGGTGGTTCGATGTAGAACGAAAGCGGCTCGAAGCTGTAAAGCGGCGTGCACAATC	420						
DB	361	TGGTGGTTCGATGTAGAACGAAAGCGGCTCGAAGCTGTAAAGCGGCGTGCACAATC	420						
QY	421	TTCTCGCCCAACGCTCAGTGGCTGATCAATTAATCTCCGCTGGATGACCAAGGATGCCA	480						
DB	421	TTCTCGCCCAACGCTCAGTGGCTGATCAATTAATCTCCGCTGGATGACCAAGGATGCCA	480						
QY	481	TTGCTGTGGAACTGCCTGCACATAATGTTCCGGCGTTATTTCTGATGTTCTGACCCAGA	540						
DB	481	TTGCTGTGGAACTGCCTGCACATAATGTTCCGGCGTTATTTCTGATGTTCTGACCCAGA	540						
QY	541	CACCCATCAACAGTATTTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600						
DB	541	CACCCATCAACAGTATTTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600						
QY	601	TGGTGCATTTGGTCAACAGCAAAATCGGCTGTTAGCGGGCCCATTAAGTTCTGTCTGG	660						
DB	601	TGGTGCATTTGGTCAACAGCAAAATCGGCTGTTAGCGGGCCCATTAAGTTCTGTCTGG	660						
QY	661	CGCGTCTGGTCTGGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
DB	661	CGCGTCTGGTCTGGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
QY	721	CGGAAACGGAAGCGCATGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA	780						
DB	721	CGGAAACGGAAGCGCATGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA	780						
QY	781	ATGAGGGCATCGTTCCCACTCGCATGCTGGTGGCCAAACGATCAGATGGCGCTGGGGCNA	840						
DB	781	ATGAGGGCATCGTTCCCACTCGCATGCTGGTGGCCAAACGATCAGATGGCGCTGGGGCNA	840						
QY	841	TGCGGCCATTACCGAGTCCGGCTGCGGTTGGTGGATATCTCGGTAGTGGGATAGC	900						
DB	841	TGCGGCCATTACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAGC	900						
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACACCATCAACAGGATTTTC	960						
DB	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACACCATCAACAGGATTTTC	960						
QY	961	GCCTGCTGGGGCAACACGAGTGGACCGCTGCTGCAACTCTCTACGGGCCAGGCGGTGA	1020						
DB	961	GCCTGCTGGGGCAACACGAGTGGACCGCTGCTGCAACTCTCTACGGGCCAGGCGGTGA	1020						

DB	961	GCCTGCTGGGGCAACACGAGCGTGACCGCTTGCTGCAACTCTCTCAGGGCCAGGGGTGA	1020						
QY	1021	AGGCAATACAGCTGTGTGCCCGTCTCAGTGTGAAAGAAAAACCACTCTGCGGCCAATA	1080						
DB	1021	AGGCAATACAGCTGTGTGCCCGTCTCAGTGTGAAAGAAAAACCACTCTGCGGCCAATA	1080						
QY	1081	CGAAAAACCGCTCTCCCGCGCGTTGGCCGATTCATTAATCGAGCTGCGACGACAGGTTT	1140						
DB	1081	CGAAAAACCGCTCTCCCGCGCGTTGGCCGATTCATTAATCGAGCTGCGACGACAGGTTT	1140						
QY	1141	CCGACTTGGAAAGCGGCGAGTGAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTTAG	1200						
DB	1141	CCGACTTGGAAAGCGGCGAGTGAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTTAG	1200						
QY	1201	GCACAAATTCATGTTTGACAGCTTATCATCGACTGACGCTGCACCAATGCTTCTGCGG	1260						
DB	1201	GCACAAATTCATGTTTGACAGCTTATCATCGACTGACGCTGCACCAATGCTTCTGCGG	1260						
QY	1261	TCAGCAGCCATCGGAAGCTGTGATGCTGTGAGGTCTGAATCACTGCATTAATTCG	1320						
DB	1261	TCAGCAGCCATCGGAAGCTGTGATGCTGTGAGGTCTGAATCACTGCATTAATTCG	1320						
QY	1321	TGTCGCTCAAGCGCACCTCCCGTCTTGATTAATGTTTTTGGCGCGACATCATACGGTT	1380						
DB	1321	TGTCGCTCAAGCGCACCTCCCGTCTTGATTAATGTTTTTGGCGCGACATCATACGGTT	1380						
QY	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTGGA	1440						
DB	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTGGA	1440						
QY	1441	ATTGTAGCGGATTAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA	1500						
DB	1441	ATTGTAGCGGATTAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA	1500						
QY	1501	GCCTTCCACCAACAGGACCAATAGATTATGAAATCTGAAGAGGTAAACTGTGTAATCTGG	1560						
DB	1501	GCCTTCCACCAACAGGACCAATAGATTATGAAATCTGAAGAGGTAAACTGTGTAATCTGG	1560						
QY	1561	ATTAACGCGATTAAGGCTATTAACGCTCTGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620						
DB	1561	ATTAACGCGATTAAGGCTATTAACGCTCTGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620						
QY	1621	ACCGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCGCCAGGTT	1680						
DB	1621	ACCGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCGCCAGGTT	1680						
QY	1681	CGGCAACTGCGGATGGCCCTGACATTAATCTCTGGGCACACGCGCTTTGGTGGCTAC	1740						
DB	1681	CGGCAACTGCGGATGGCCCTGACATTAATCTCTGGGCACACGCGCTTTGGTGGCTAC	1740						
QY	1741	GCTCAATCTGGCTGTTGGCTGAAATCAACCCGGACAAAGCGTTCCAGGACAAGCTGTAT	1800						
DB	1741	GCTCAATCTGGCTGTTGGCTGAAATCAACCCGGACAAAGCGTTCCAGGACAAGCTGTAT	1800						
QY	1801	CGTTTACCTGGATGCGGTACGTTTACAAACGCAAGCTGATGTTACCGATCCCTGTT	1860						
DB	1801	CGTTTACCTGGATGCGGTACGTTTACAAACGCAAGCTGATGTTACCGATCCCTGTT	1860						
QY	1861	GAAGGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA	1920						
DB	1861	GAAGGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA	1920						
QY	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGGTGAAGAGCGCGCTGATGTTCAAC	1980						
DB	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGGTGAAGAGCGCGCTGATGTTCAAC	1980						
QY	1981	CTGCAAGAACCGTACTTCACTGGCCGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG	2040						
DB	1981	CTGCAAGAACCGTACTTCACTGGCCGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG	2040						
QY	2041	TATGAAAACGCGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGGAAAGCG	2100						
DB	2041	TATGAAAACGCGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGGAAAGCG	2100						


```
QY 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAACAAACACATGAATCGACACCCGATTAC 2160
DB 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAACAAACACATGAATCGACACCCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAATAAGGCGAAACAGCGATGACCATCAACGCGCCGCTGG 2220
DB 2161 TCCATCGCAGAGCTGCTTTTAATAAGGCGAAACAGCGATGACCATCAACGCGCCGCTGG 2220
QY 2221 GCATGGTCCAAATCGACACCCAGCAAGTGAATTATGGTGAACGCTACTGCGGACCTTC 2280
DB 2221 GCATGGTCCAAATCGACACCCAGCAAGTGAATTATGGTGAACGCTACTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTCTGGTGGCTGCTGAGCGCAGGTATTAAACGCCGCCAGT 2340
DB 2281 AAGGGTCAACCATCCAAACCGTTCTGGTGGCTGCTGAGCGCAGGTATTAAACGCCGCCAGT 2340
QY 2341 CGAACAAAGAGCTGCGCAAGAGTTCTCGAAACATATCTGCTGACTGATGAAGTCTG 2400
DB 2341 CGAACAAAGAGCTGCGCAAGAGTTCTCGAAACATATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
DB 2401 GAAGCGGTTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAGATCCAGTATTGCGGCCACCATGGAACACGCCAGAAAGGTGAATCATG 2520
DB 2461 TTGGCGAAGATCCAGTATTGCGGCCACCATGGAACACGCCAGAAAGGTGAATCATG 2520
QY 2521 CGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAAGCC 2580
DB 2521 CGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAAGCC 2580
QY 2581 GCCAGCGGTGTCAGATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
DB 2581 GCCAGCGGTGTCAGATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAACAATAACAACAACCTCGGATCGAGGGAAGATTCAGAATTC 2700
DB 2641 AACAAACAACAATAACAATAACAACAACCTCGGATCGAGGGAAGATTCAGAATTC 2700

RESULT 14
US-10-263-153-20
; Sequence 20, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Gineburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263.153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 7370
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3447)
; OTHER INFORMATION: pMBP-c2X-Toxo30del13C (52-300aa)
US-10-263-153-20
```

Query Match 81.5%; Score 2690.4; DB 19; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
QY 1 CCGACACCATCGAATGGTGTCAAAACCTTTCCGGTATGGCATGATAGCGCCCGAAGAGA 60
DB 1 CCGACACCATCGAATGGTGTCAAAACCTTTCCGGTATGGCATGATAGCGCCCGAAGAGA 60
QY 61 GTCAAATTCAGAGGTGGTGAATGTGAACCAAGTAACTTTATAGATGTCGAGAGTATGCGG 120
DB 61 GTCAAATTCAGAGGTGGTGAATGTGAACCAAGTAACTTTATAGATGTCGAGAGTATGCGG 120
QY 121 GTGTCTCTTATCAGACCGCTTTCCCGCGTGGTGAACCAAGCGAGCCAGCTTTCTCGGAAA 180
DB 121 GTGTCTCTTATCAGACCGCTTTCCCGCGTGGTGAACCAAGCGAGCCAGCTTTCTCGGAAA 180
QY 181 CGCGGGAAGAGTGAAGCGCGGATGCGGAGCTGAATTAATTTCCCAACCGGTGGCAC 240
DB 181 CGCGGGAAGAGTGAAGCGCGGATGCGGAGCTGAATTAATTTCCCAACCGGTGGCAC 240
QY 241 AACAACTGGCGGGGCAACAGTCTGCTGATTCGGCGTTGCCAGCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGGGCAACAGTCTGCTGATTCGGCGTTGCCAGCTCCAGTCTGGCCCTGC 300
QY 301 ACGCCCGTGCAGAAATTTGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
DB 301 ACGCCCGTGCAGAAATTTGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTGTGATGAGTGAAGCGCGGTGCAAGCTGTAAAGCGCGGTGCACAATC 420
DB 361 TGGTGGTGTGATGAGTGAAGCGCGGTGCAAGCTGTAAAGCGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTCGCGCGATCAACTGGGTGCCAG 480
DB 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTCGCGCGATCAACTGGGTGCCAG 480
QY 481 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTCGCGCGATCAACTGGGTGCCAG 540
DB 481 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTCGCGCGATCAACTGGGTGCCAG 540
QY 541 CACCCATCAACAGTATTTCTCCATGAAGCGGTAGCGACTGGCGGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTTCTCCATGAAGCGGTAGCGACTGGCGGTGGAGCATC 600
QY 601 TGGTGGTGTGATGAGTGAAGCGCGGTGCAAGCTGTAAAGCGCGGTGCACAATC 660
DB 601 TGGTGGTGTGATGAGTGAAGCGCGGTGCAAGCTGTAAAGCGCGGTGCACAATC 660
QY 661 CCGCTCTGCTGCTGCGTGGCTGCGATTAATTAATCTCACTCGGCAATCAATTCAGCGCAT 720
DB 661 CCGCTCTGCTGCTGCGTGGCTGCGATTAATTAATCTCACTCGGCAATCAATTCAGCGCAT 720
QY 721 CGGAACGGGAAGCGGACTGGAGTGCATGTCGGGTTCACAAACCATGCAAAATGCTGA 780
DB 721 CGGAACGGGAAGCGGACTGGAGTGCATGTCGGGTTCACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCTTCCCATGCGATGCTGGTTGCCAAGATGCGGTGGCGCGCAA 840
DB 781 ATGAGGGCATCTTCCCATGCGATGCTGGTTGCCAAGATGCGGTGGCGCGCAA 840
QY 841 TGGCGGCCATTAACCGAGTCCGGCTGCGGTGGTGGCGGATATCTCGGTAGTGGATAG 900
DB 841 TGGCGGCCATTAACCGAGTCCGGCTGCGGTGGTGGCGGATATCTCGGTAGTGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCATCAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCATCAACAGGATTTTC 960
QY 961 GCGTCTGGGCAACACAGCGTGGACCGCTTGGTGGCGGATCTCTCAGGGCCAGGGGTGA 1020
DB 961 GCGTCTGGGCAACACAGCGTGGACCGCTTGGTGGCGGATCTCTCAGGGCCAGGGGTGA 1020
QY 1021 AGGCAATCAGTGTGCGGTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
DB 1021 AGGCAATCAGTGTGCGGTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
QY 1081 CGCAACCCGCTCTCCCGCGGTTGGCGGATTCATTAATGACGTGCGACGACAGGTTT 1140
```

Db 1081 |||||CGCAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAGACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTCTACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTCTACTCATTTAG 1200
Qy 1201 GCACAATTTCTCATGTTTGACAGCTTATCATGACCTGCGACGCGTGCACCAATGCTTCGCG 1260
Db 1201 GCACAATTTCTCATGTTTGACAGCTTATCATGACCTGCGACGCGTGCACCAATGCTTCGCG 1260
Qy 1261 TCAGCAGCCATCGGAAGCTGCGTATGCGTGTGAGGTGCGTAAATCACTGCATAAATTCG 1320
Db 1261 TCAGCAGCCATCGGAAGCTGCGTATGCGTGTGAGGTGCGTAAATCACTGCATAAATTCG 1320
Qy 1321 TGTCTCTCAAGCGGCACCTCCCGTCTCGGATATGTTTTTGGCGCGACATCAATACGGTT 1380
Db 1321 TGTCTCTCAAGCGGCACCTCCCGTCTCGGATATGTTTTTGGCGCGACATCAATACGGTT 1380
Qy 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATTCG 1440
Db 1381 CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATTCG 1440
Qy 1441 ATTGTGAGCGGATACAAATTTACACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Db 1441 ATTGTGAGCGGATACAAATTTACACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Qy 1501 GCACCTTACCAACAGGACCATAGATTTATGAAAACTGGAAGAGTTAACTGGTAAATTCG 1560
Db 1501 GCACCTTACCAACAGGACCATAGATTTATGAAAACTGGAAGAGTTAACTGGTAAATTCG 1560
Qy 1561 ATTAACCGCGATAAAGGCTATTAACCGTCTCGTGAAGTTCGTTAGTGTTCACGA 1620
Db 1561 ATTAACCGCGATAAAGGCTATTAACCGTCTCGTGAAGTTCGTTAGTGTTCACGA 1620
Qy 1621 ACCGGAATTAAGTACCGTTGAGCATCCGATTAATTCGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTACCGTTGAGCATCCGATTAATTCGGAAGAGAAATTCACAGGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTGGGACACAGCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTGGGACACAGCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCGGCAAGCGTTCCAGGACAAAGCTGAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCGGCAAGCGTTCCAGGACAAAGCTGAT 1800
Qy 1801 CCGTTTACCTGGATGCGGTACGTTTACAGCGCAAGCTGATTCCTTACCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGATGCGGTACGTTTACAGCGCAAGCTGATTCCTTACCGATCGCTGT 1860
Qy 1861 GAAGGTTTATCGCTGATTTATACAAAGATCTGCTGCGAACCCTGCAAAACCTGGGAA 1920
Db 1861 GAAGGTTTATCGCTGATTTATACAAAGATCTGCTGCGAACCCTGCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGTGAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGTGAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGTTGCTGACGGGGTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGTTGCTGACGGGGTTATCGGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTAAGACGTTGGCGTGAATGCGTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGTTGGCGTGAATGCGTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGTTGACCTGATTTAAAAACAAACATGATGACAGACCGGATAC 2160
Db 2101 GGTCTGACCTTCTGTTGACCTGATTTAAAAACAAACATGATGACAGACCGGATAC 2160
Qy 2161 TCCATTCGCAAGAGCTGCCTTTTAAAGCGGAAACAGCGATGACCATCAACCGGCGGTG 2220

Db 2161 TCCATTCGCAAGAGCTGCCTTTTAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCCAAACATCGACACAGCAAAAGTGAATTTATGGTTAAACGGTACTCCGACCTTC 2280
Db 2221 GCATGGTCCAAACATCGACACAGCAAAAGTGAATTTATGGTTAAACGGTACTCCGACCTTC 2280
Qy 2281 AAGGTTCAACCATCCAAACCGTTTGGTGGCGTCTGAGCGCAGGTATTTAAACGGCCGAGT 2340
Db 2281 AAGGTTCAACCATCCAAACCGTTTGGTGGCGTCTGAGCGCAGGTATTTAAACGGCCGAGT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCCGAAACATCTCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCCGAAACATCTCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAAGCGGTTAATAAGACAAACCCGCTGGTCCGCTAGCGTGAAGTCTTTACGAGGAAG 2460
Db 2401 GAAAGCGGTTAATAAGACAAACCCGCTGGTCCGCTAGCGTGAAGTCTTTACGAGGAAG 2460
Qy 2461 TTGGGAAAGATCCACGTATTTGGCGCCACATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGGAAAGATCCACGTATTTGGCGCCACATGGAACCGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGGTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGGTGCTGCTGCTGCTGCTGCTGCTG 2580
Qy 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGGAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGGAAGACGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACAAACAAACATTAACAAACAAACAAACCTCGGGATCGAGGAGGATTCAGAAATTC 2700
Db 2641 AACAAACAAACATTAACAAACAAACAAACCTCGGGATCGAGGAGGATTCAGAAATTC 2700

RESULT 15

US-10-153-61
; Sequence 61, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 7370
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3447)
; OTHER INFORMATION: pMBP-c2X-Toxop30NIX1
US-10-263-153-61

Query Match 81.5%; Score 2690.4; DB 18; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 CCGACCATCGAATGGTGC AAAACCTTTTCGCGGTATCGCATGATAGCGCCCGGAAGAGA 60
Db 1 CCGACCATCGAATGGTGC AAAACCTTTTCGCGGTATCGCATGATAGCGCCCGGAAGAGA 60
Qy 61 GTCATTCAGGTGTTGATGTGAACCAAGTAACTTATACGATTCGAGAGTATGCG 120
Db 61 GTCATTCAGGTGTTGATGTGAACCAAGTAACTTATACGATTCGAGAGTATGCG 120

QY	121	GTGCTCTTATCAGACCGGTTTCCGGGTGGTGAACAGGCCAGCCACGCTTTCTGCGAAA	180
Db	121	GTGCTCTTATCAGACCGGTTTCCGGGTGGTGAACAGGCCAGCCACGCTTTCTGCGAAA	180
QY	181	CGCGGGAAGAGTGGAGCGCGGATGCGGAGCTGAATTAATTCACATTCACACCGCGTGGCAC	240
Db	181	CGCGGGAAGAGTGGAGCGCGGATGCGGAGCTGAATTAATTCACATTCACACCGCGTGGCAC	240
QY	241	AACAACTGGCGGCAAAACAGTCTGCTGATTTGGCGTTCGCACCTCCAGTCTGGCGCTGC	300
Db	241	AACAACTGGCGGCAAAACAGTCTGCTGATTTGGCGTTCGCACCTCCAGTCTGGCGCTGC	300
QY	301	ACGCGCGTGGCAAAATGTCGCGCGATTAATCTCGCGCGGATCAATCGGCTGCCAGCG	360
Db	301	ACGCGCGTGGCAAAATGTCGCGCGATTAATCTCGCGCGGATCAATCGGCTGCCAGCG	360
QY	361	TGCTGTGTCGATGGTAGAACGACGCGCTGAAAGCTGTAAAGCGCGGTGCAACAATC	420
Db	361	TGCTGTGTCGATGGTAGAACGACGCGCTGAAAGCTGTAAAGCGCGGTGCAACAATC	420
QY	421	TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCCA	480
Db	421	TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCCA	480
QY	481	TTGCTGTGGAAGCTGCTGCACTAATGTTTCGGCGTTATTTCTTGATGTTCTCTGACCCAGA	540
Db	481	TTGCTGTGGAAGCTGCTGCACTAATGTTTCGGCGTTATTTCTTGATGTTCTCTGACCCAGA	540
QY	541	CACCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACGGACTGGCGGTGGAGCATC	600
Db	541	CACCCATCAACAGTATTATTTTCTCCCATGAAGCGGTACGGACTGGCGGTGGAGCATC	600
QY	601	TGGTGCATTTGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAAGTTCTGTCTCGG	660
Db	601	TGGTGCATTTGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAAGTTCTGTCTCGG	660
QY	661	CGCGTCTGGCTGGCTGGGTAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Db	661	CGCGTCTGGCTGGCTGGGTAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGAAACGGGAAGGACATGGAGTGCCATGTCGGGTTTTCAACAACCATGCAAAATGCTGA	780
Db	721	CGAAACGGGAAGGACATGGAGTGCCATGTCGGGTTTTCAACAACCATGCAAAATGCTGA	780
QY	781	ATGAGGGCATCTGTTCCCATCGGATGCTGTTGCAACAGATCAGATGGCGTGGCGCAA	840
Db	781	ATGAGGGCATCTGTTCCCATCGGATGCTGTTGCAACAGATCAGATGGCGTGGCGCAA	840
QY	841	TGGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG	900
Db	841	TGGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACGGAAGACAGCTCATGTTTATATCCCGCGTTAAACCCATCAAAACAGGATTTTC	960
Db	901	ACGATACGGAAGACAGCTCATGTTTATATCCCGCGTTAAACCCATCAAAACAGGATTTTC	960
QY	961	GCCTGCTGGGGCAACAGCGTGACCGCTTCTGCAACTCTCTCAGGGCGACGCGTGA	1020
Db	961	GCCTGCTGGGGCAACAGCGTGACCGCTTCTGCAACTCTCTCAGGGCGACGCGTGA	1020
QY	1021	AGGCAATACAGCTGTTGCGGCTCTCACTGTGTAAGAAAACACCCCTGGCGCCCAATA	1080
Db	1021	AGGCAATACAGCTGTTGCGGCTCTCACTGTGTAAGAAAACACCCCTGGCGCCCAATA	1080
QY	1081	CGMAACCGGCTCTCCCGCGGTTGGCGATTCAATTAATGCAAGTGGCACGACAGGTTT	1140
Db	1081	CGMAACCGGCTCTCCCGCGGTTGGCGATTCAATTAATGCAAGTGGCACGACAGGTTT	1140
QY	1141	CCGACTGGAAAGCGGCGAGTGAACGCAATTAATGTTAGTTAGTCACTCATTTAG	1200
Db	1141	CCGACTGGAAAGCGGCGAGTGAACGCAATTAATGTTAGTTAGTCACTCATTTAG	1200

QY	1201	GCACAAATTCATGTTTGACAGCTTATCATCGACTGCACCGTGCAACCAATCTCTCTGGCG	1260
Db	1201	GCACAAATTCATGTTTGACAGCTTATCATCGACTGCACCGTGCAACCAATCTCTCTGGCG	1260
QY	1261	TCAGGACGCCAATCGGAAGCTGTGATGGCTGTGTCAGGTCGTAAATCACTGCATAAATTCG	1320
Db	1261	TCAGGACGCCAATCGGAAGCTGTGATGGCTGTGTCAGGTCGTAAATCACTGCATAAATTCG	1320
QY	1321	TGTCGCTCAAGCGCACCTCCCGTTCTCGATAATGTTTTTTTGGCGCGACATCATAAACGGTT	1380
Db	1321	TGTCGCTCAAGCGCACCTCCCGTTCTCGATAATGTTTTTTTGGCGCGACATCATAAACGGTT	1380
QY	1381	CTGCAAAATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA	1440
Db	1381	CTGCAAAATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA	1440
QY	1441	ATTGTAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAAAGTGTGTTTACGA	1500
Db	1441	ATTGTAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAAAGTGTGTTTACGA	1500
QY	1501	GCATTCACCAACAGGACCATAGATTATCAAAAATGGAAGGTAAACTGGTAATCTGG	1560
Db	1501	GCATTCACCAACAGGACCATAGATTATCAAAAATGGAAGGTAAACTGGTAATCTGG	1560
QY	1561	ATTAAACGGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGGAAGAAT	1620
Db	1561	ATTAAACGGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGGAAGAAT	1620
QY	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCACAGGTT	1680
Db	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCACAGGTT	1680
QY	1681	GGCGCACTGGCGATGGCCCTGACATTAATCTCTGGGCAACGACCGCTTGGTGCTAC	1740
Db	1681	GGCGCACTGGCGATGGCCCTGACATTAATCTCTGGGCAACGACCGCTTGGTGCTAC	1740
QY	1741	GCTCAATCTGGCCCTGTTGGCTGAATTCACCCGGACAAAGCGTTCCAGACAGCTGTAT	1800
Db	1741	GCTCAATCTGGCCCTGTTGGCTGAATTCACCCGGACAAAGCGTTCCAGACAGCTGTAT	1800
QY	1801	CGGTTTACCTGGGATGCGGTGATTAACAACGGCAAGCTGATTCCTACCCGATCGCTGTT	1860
Db	1801	CGGTTTACCTGGGATGCGGTGATTAACAACGGCAAGCTGATTCCTACCCGATCGCTGTT	1860
QY	1861	GAAGCGTTATCGCTGATTTTATACAAAGATCTGCTGCCAAACCCGCCAAAACCTGGGAA	1920
Db	1861	GAAGCGTTATCGCTGATTTTATACAAAGATCTGCTGCCAAACCCGCCAAAACCTGGGAA	1920
QY	1921	GAGATCCCGCGCTGGATAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTACCTGGCGCTGATGCTGTGACGGGGTTATCGGTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTACCTGGCGCTGATGCTGTGACGGGGTTATCGGTTCAAG	2040
QY	2041	TATGAAAACGGCAAGTACGACATTAAGACGTTGAGCGTGAATAACGCTGGCGCGAAGCG	2100
Db	2041	TATGAAAACGGCAAGTACGACATTAAGACGTTGAGCGTGAATAACGCTGGCGCGAAGCG	2100
QY	2101	GCTCTGAACCTTCTCGTGGTGAACCTGATTAATAACACATGAATGCAGACACCGATTAC	2160
Db	2101	GCTCTGAACCTTCTCGTGGTGAACCTGATTAATAACACATGAATGCAGACACCGATTAC	2160
QY	2161	TCATTCGAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACCGGCCGTTG	2220
Db	2161	TCATTCGAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACCGGCCGTTG	2220
QY	2221	GCATGGTCCAAATCGACACCGCAAGTGAATTAATGTTGTAACCGTACTGCGACCTTC	2280
Db	2221	GCATGGTCCAAATCGACACCGCAAGTGAATTAATGTTGTAACCGTACTGCGACCTTC	2280
QY	2281	AGGGTCAACCATCCAAACCGTTCTGGCGGTGCTGAGCGCAGGTATTAAACGCGCCAGT	2340

[illegible]

Search completed: August 3, 2005, 04:21:42
Job time : 1322.2 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 14:24:04 ; Search time 1094.6 Seconds
(without alignments)
17846.836 Million cell updates/sec

Title: US-09-765-555B-14
Perfect score: 3300
Sequence: 1 ccgcacacatgaatgtgc.....acgacgttcaggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3300	100.0	3300	4 AAD11588	Aad11588 Partial s
2	3300	100.0	3300	4 AAD11589	Aad11589 Partial s
3	3226.4	97.8	3300	4 AAD11592	Aad11592 Partial s
4	3215.2	97.4	3300	4 AAD11591	Aad11591 Partial s
5	3207.2	97.2	3300	4 AAD11590	Aad11590 Partial s
6	2705.8	82.0	8101	4 AAD25681	Aad25681 Nucleotid
7	2705.8	82.0	8101	6 ABL53238	ABL53238 Nucleotid
8	2690.4	81.5	7259	12 ADO23608	Ado23608 DNA encod
9	2690.4	81.5	7322	12 ADO23613	Ado23613 DNA encod
10	2690.4	81.5	7352	12 ADO23603	Ado23603 DNA encod
11	2690.4	81.5	7370	12 ADO23598	Ado23598 DNA encod
12	2690.4	81.5	7370	12 ADO23639	Ado23639 DNA encod
13	2690.4	81.5	7370	12 ADO23649	Ado23649 DNA encod
14	2690.4	81.5	7370	12 ADO23644	Ado23644 DNA encod
15	2690.4	81.5	7403	12 ADO23594	Ado23594 DNA encod
16	2690.4	81.5	7442	12 ADO23590	Ado23590 DNA encod
17	2690.4	81.5	7478	12 ADO23584	Ado23584 DNA encod
18	2688.8	81.5	7112	12 ADO23618	Ado23618 DNA encod
19	2636.8	79.9	6806	6 ABL49925	ABL49925 Maltose b
20	2605.4	79.0	7553	12 ADO23588	Ado23588 DNA encod

21	1609.4	48.8	5558	12	ADL90419	Adl90419 Clostridi
22	1474	44.7	4700	9	ACF06053	Actf06053 Vector pJ
23	1472.4	44.6	4700	12	ADL18670	Adl18670 Vector pJ
24	1462.4	44.3	4920	10	ADH73599	Adh73599 Plasmid p
25	1462.4	44.3	4935	10	ADH73598	Adh73598 Plasmid p
26	1462.4	44.3	4945	10	ADH73600	Adh73600 Plasmid p
27	1462.4	44.3	4951	10	ADH73601	Adh73601 Plasmid p
28	1418.2	43.0	5903	8	ABZ23939	Abz23939 Nucleotid
29	1417.6	43.0	1511	6	ABQ73200	Abq73200 E. coli t
30	1417.6	43.0	5926	2	AAV32977	Aav32977 Tn7 donor
31	1417.6	43.0	5926	6	AAD45059	Aad45059 Transposo
32	1417.6	43.0	5926	12	ADG46817	Adg46817 Donor pla
33	1362	41.3	5201	12	ADL72229	Adl72229 DNA seque
34	1362	41.3	5201	12	ADL72228	Adl72228 DNA seque
35	1238.4	37.5	5024	9	ACF06051	Actf06051 Plasmid p
36	1238.4	37.5	5024	12	ADL18668	Adl18668 Plasmid p
37	1237.2	37.5	1922	6	ABQ73089	Abq73089 MBP and z
38	1237.2	37.5	1922	10	ADF83622	Adf83622 MBP-human
39	1204	36.5	4557	2	AAT90491	Aat90491 Vector pl
40	1204	36.5	4969	12	ADO33522	Ado33522 Expressio
41	1204	36.5	5006	10	ADE24112	Ade24112 plasmid p
42	1204	36.5	5064	12	ADO33532	Ado33532 Vector pG
43	1204	36.5	5082	12	ADO33531	Ado33531 Vector pG
44	1204	36.5	6259	6	AAZ29720	Aaz29720 Plasmid p
45	1204	36.5	6823	3	AAC55459	Aac55459 Destinati

ALIGNMENTS

RESULT 1	
AAAD11588	
ID	AAAD11588 standard; DNA; 3300 BP.
XX	
AC	AAAD11588;
XX	
DT	24-SEP-2001 (first entry)
XX	
DE	Partial sequence of pMal-m1 and ZFPm1 DNA.
XX	
KW	Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
KW	modulation; plant technology; agriculture; ds.
XX	
OS	Unidentified.
XX	
PH	Key
FT	CDS
FT	Location/Qualifiers
FT	2719..3270
FT	/tag= a
FT	/product= "ZFPm1 protein"
FT	/transl_except= (pos:2785..2787, aa:Asp)
FT	/transl_except= (pos:2788..2790, aa:Pro)
FT	/transl_except= (pos:2791..2793, aa:Gly)
FT	/transl_except= (pos:2872..2874, aa:Arg)
FT	/transl_except= (pos:2875..2877, aa:Ala)
FT	/transl_except= (pos:2878..2880, aa:His)
FT	/transl_except= (pos:2884..2886, aa:Glu)
FT	/transl_except= (pos:2953..2955, aa:Gln)
FT	/transl_except= (pos:2959..2961, aa:Ser)
FT	/transl_except= (pos:3121..3123, aa:Arg)
FT	/transl_except= (pos:3124..3126, aa:Ser)
FT	/transl_except= (pos:3127..3129, aa:Asp)
FT	/transl_except= (pos:3130..3132, aa:Asn)
FT	/transl_except= (pos:3136..3138, aa:Val)
FT	/transl_except= (pos:3139..3141, aa:Arg)
FT	/transl_except= (pos:3205..3207, aa:Ala)
FT	/transl_except= (pos:3208..3210, aa:Ala)
FT	/transl_except= (pos:3211..3213, aa:Gly)
FT	/transl_except= (pos:3214..3216, aa:His)
FT	/transl_except= (pos:3220..3222, aa:Ala)
FT	/transl_except= (pos:3223..3225, aa:Ser)
FT	/note= "CDS does not include start and stop codon"
FT	/partial
FT	primer_bind
FT	2740..2790


```

FT primer_bind /tag= b
FT /bound_moiety= "F1-f2 primer"
FT 2770.._2850
FT /tag= c
FT /bound_moiety= "F1-f1 primer"
FT 2824.._2889
FT /tag= d
FT /bound_moiety= "F2-b primer"
FT 2867.._2940
FT /tag= e
FT /bound_moiety= "F2-f primer"
FT 2916.._2973
FT /tag= f
FT /bound_moiety= "F3-b1 primer"
FT 2953.._3021
FT /tag= g
FT /bound_moiety= "F3-b2 primer"
FT 2992.._3042
FT /tag= h
FT /bound_moiety= "F4-f2 primer"
FT 3022.._3102
FT /tag= i
FT /bound_moiety= "F4-f1 primer"
FT 3076.._3141
FT /tag= j
FT /bound_moiety= "F5-b primer"
FT 3119.._3192
FT /tag= k
FT /bound_moiety= "F5-f primer"
FT 3168.._3225
FT /tag= l
FT /bound_moiety= "F6-b1 primer"
FT 3205.._3273
FT /tag= m
FT /bound_moiety= "F6-b2 primer"

```

WO200152620-A2.

26-JUL-2001.

19-JAN-2001; 2001WO-US001817.

21-JAN-2000; 2000US-0177468P.

21-JUL-2000; 2000US-00620897.

(SCRI) SCRIPPS RES INST.

(SYGN) SYNGENTA AGRIC DISCOVERY INC.

Barbas CF, Stege JT, Guan X, Dalmia B;

WPI; 2001-465325/50.

P-PSDB; AAE06000.

New zinc finger proteins, useful for modulating or regulating gene expression and metabolic pathways in plants, e.g. for treating in the plant cells a disorder that is associated with abnormal expression of the target gene.

Example 4; Page 138-139; 156pp; English.

The patent discloses methods and compositions to modulate the expression of a target gene in plant cells. The method involves providing plant cells with a zinc finger protein (ZFP) which is capable of specifically binding to a target nucleotide sequence or its complementary strand within a target gene and allowing the ZFP binding to the target nucleotide sequence, where the expression of the target gene in the plant cells is modulated. The ZFP and fusions of the ZFP proteins are useful for modulating or regulating gene expression and metabolic pathways in plants. The ZFP, fusion proteins and methods are useful in plant and agricultural technology. The method is useful particularly for treating a disorder in the plant cells, where the disorder is associated with abnormal expression of the target gene. The present DNA sequence is the partial sequence of pMal-m1 and ZFPm1 DNA

```

XX SQ Sequence 3300 BP; 847 A; 874 C; 876 G; 703 T; 0 U; 0 Other;
Query Match 100.0%; Score 3300; DB 4; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGTGCAAAACCTTTTCGGGTATGGCATGATAGCGCCGGAAGAGA 60
DB 1 CCGACACCATCGAATGGTGTGCAAAACCTTTTCGGGTATGGCATGATAGCGCCGGAAGAGA 60
QY 61 GTCAAATTCAGGGTGGTGAATGTGAAACCAAGTAAGCTTATAGATGTCGAGAGTATCCG 120
DB 61 GTCAAATTCAGGGTGGTGAATGTGAAACCAAGTAAGCTTATAGATGTCGAGAGTATCCG 120
QY 121 GTGTCTCTTATCAGACCCGTTTCCCGCGTGTGTAACACGAGCCAGCCAGCTTCTCGAAAA 180
DB 121 GTGTCTCTTATCAGACCCGTTTCCCGCGTGTGTAACACGAGCCAGCCAGCTTCTCGAAAA 180
QY 181 CGCGGGAAGAGTGAAGCGCGCATGGGAGCTGAATTACATTCCTCCAAACGGGTGGAC 240
DB 181 CGCGGGAAGAGTGAAGCGCGCATGGGAGCTGAATTACATTCCTCCAAACGGGTGGAC 240
QY 241 AACAACTGGCGGCAAAACAGTTCGTGTGATTTGGCGTTGCCACCTCCAGTCTGSCCTGC 300
DB 241 AACAACTGGCGGCAAAACAGTTCGTGTGATTTGGCGTTGCCACCTCCAGTCTGSCCTGC 300
QY 301 ACGCGCGTCGCAAAATTTGTCCGCGCATTAATCTCGCGCGATCAACTCGGTGCCAGCG 360
DB 301 ACGCGCGTCGCAAAATTTGTCCGCGCATTAATCTCGCGCGATCAACTCGGTGCCAGCG 360
QY 361 TGGTGGTGTGATAGAGAGCGGCGTGAAGCTGTAAAGCGCGGTGACAAATC 420
DB 361 TGGTGGTGTGATAGAGAGCGGCGTGAAGCTGTAAAGCGCGGTGACAAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATTAATTCGCTGGATGACAGATGCCA 480
DB 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATTAATTCGCTGGATGACAGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACATAATTTTCGGCGGTATTTCTTGATGTCTTGACAGA 540
DB 481 TTGCTGTGGAAGCTGCTGCACATAATTTTCGGCGGTATTTCTTGATGTCTTGACAGA 540
QY 541 CACCATCAACAGTATTATTTTCTCCCATGAAGACGCTAGCGACTGGGGGTGGAGATC 600
DB 541 CACCATCAACAGTATTATTTTCTCCCATGAAGACGCTAGCGACTGGGGGTGGAGATC 600
QY 601 TGGTGGATTTGGTCCACAGCAAAATCGCGTGTAGCGGCGCCATTAAGTTCTGTCTCG 660
DB 601 TGGTGGATTTGGTCCACAGCAAAATCGCGTGTAGCGGCGCCATTAAGTTCTGTCTCG 660
QY 661 CGGCTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCGATAG 720
DB 661 CGGCTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCGATAG 720
QY 721 CGGAAACGGGAAGCGGACTGGAGTGCCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
DB 721 CGGAAACGGGAAGCGGACTGGAGTGCCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGCAATCGTTTCCACTCGCATCTGTTGCCAACGATCAGATGGCGCTGGGGCGAA 840
DB 781 ATGAGGCAATCGTTTCCACTCGCATCTGTTGCCAACGATCAGATGGCGCTGGGGCGAA 840
QY 841 TGGCGGCCCATTCAGGAGTCCGGCTGGCGTGGTGGGATATCTCGGTAGTGGGATACG 900
DB 841 TGGCGGCCCATTCAGGAGTCCGGCTGGCGTGGTGGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACCAACCATCAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACCAACCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGGCAACACAGGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGGCGTGA 1020
DB 961 GCCTGCTGGGGCAACACAGGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGGCGTGA 1020

```

Db 961 GCCTGCTGGGCAAAACCAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAAGAAAAACCAACCCCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAAGAAAAACCAACCCCTGGCGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGACAGTGGCGACGACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGACAGTGGCGACGACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGGCAGTGAGGCGCAACGCAATTAATGTGAGTTAGTCACTCAATTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGAGGCGCAACGCAATTAATGTGAGTTAGTCACTCAATTAG 1200
Qy 1201 GCACAATCTCATGTTTGAAGCTTATCATCGACTGCAAGCGGTGCAACCAATGCTTCTGGCG 1260
Db 1201 GCACAATCTCATGTTTGAAGCTTATCATCGACTGCAAGCGGTGCAACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGAGCGCATCGGAAGCTGGTATGCTGTGACAGTTCGTAATCACTGATCAATTCG 1320
Db 1261 TCAGGAGCGCATCGGAAGCTGGTATGCTGTGACAGTTCGTAATCACTGATCAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTCTTGGATAATGTTTTTGGCGGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTTGGATAATGTTTTTGGCGGACATCAATACGGTT 1380
Qy 1381 CTGGCAATATCTGAATGAGCTGTTGTAATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAATATCTGAATGAGCTGTTGTAATTAATCATCGGCTCGTATATGTTGGA 1440
Qy 1441 ATTGTGAGCGGATACAAATTTCAACAGGAACAGCCAGTCCGTTTAGTGTTTTTCAGA 1500
Db 1441 ATTGTGAGCGGATACAAATTTCAACAGGAACAGCCAGTCCGTTTAGTGTTTTTCAGA 1500
Qy 1501 GCACCTTCAACCAACAGGACATAGATTATGAAAACTGAAAGAGTTAAACTGTTAATCTGG 1560
Db 1501 GCACCTTCAACCAACAGGACATAGATTATGAAAACTGAAAGAGTTAAACTGTTAATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAAGCGTCTCGTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAAGCGTCTCGTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACAGGTT 1680
Qy 1681 GCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGCGCTTCTGCTGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGCGCTTCTGCTGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGATCCGTAAGTTTCAACGCGCAAGCTGATTCCTTACCGGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATCCGTAAGTTTCAACGCGCAAGCTGATTCCTTACCGGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGTGAATTTATACAAAGATCTGCTGCGCAACCGCCAAABACCTGGGAA 1920
Db 1861 GAAGCGTTATCGTGAATTTATACAAAGATCTGCTGCGCAACCGCCAAABACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGGCTGATTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGGCTGATTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGGCCGCTGATTTGCTGTAAGAGCGGCTGATTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCCGCTGATTTGCTGTAAGAGCGGCTGATTTCAAG 2040
Qy 2041 TATGAAACCGGCAAGTACGACATTAAGACGTTGGCGTGGATAACGCTGGCGGAAGCG 2100
Db 2041 TATGAAACCGGCAAGTACGACATTAAGACGTTGGCGTGGATAACGCTGGCGGAAGCG 2100

Qy 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAAACAAACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAAACAAACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGCAACAGCGATGACCATCAAGCGGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGCAACAGCGATGACCATCAAGCGGCCGCTGG 2220
Qy 2221 GCATGGTCAACATCGACACAGCAAAAGTGAATTAATGGTTAAACGTTACTGCGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACAGCAAAAGTGAATTAATGGTTAAACGTTACTGCGACCTTC 2280
Qy 2281 AAGGTTCAACCATCAACACCGTTGTTGGCGTGTGAGCGCAGGATTAATAAGCGGCCAGT 2340
Db 2281 AAGGTTCAACCATCAACACCGTTGTTGGCGTGTGAGCGCAGGATTAATAAGCGGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGCTAGCGCTGAAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGCTAGCGCTGAAAGTCTTACGAGGAAG 2460
Qy 2461 TTGGGAAAGATTCACGTTATTTGCCGCCACCATGMAAACGCCACAGAAAGGTGAATCATG 2520
Db 2461 TTGGGAAAGATTCACGTTATTTGCCGCCACCATGMAAACGCCACAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGCTCGCTTTCTGTTGATGCGCTGCTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGCTCGCTTTCTGTTGATGCGCTGCTGCTGCTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTG 2640
Db 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTG 2640
Qy 2641 AACAAACAAACATAAACAAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAAACATAAACAAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Qy 2701 GGATCCTCTCTCTGTCGCGCCAGCGCGCTCGAGCGCGGGGAGAACCTTATCTTGT 2760
Db 2701 GGATCCTCTCTCTGTCGCGCCAGCGCGCTCGAGCGCGGGGAGAACCTTATCTTGT 2760
Qy 2761 CCGGAATGTGGTAAGTCTCTCTCAGAGCTCTCACCTGCTGCGCCACAGCGTACCAC 2820
Db 2761 CCGGAATGTGGTAAGTCTCTCTCAGAGCTCTCACCTGCTGCGCCACAGCGTACCAC 2820
Qy 2821 ACGGTTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGTTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTTAGCCAGTCCAGCAAC 2880
Qy 2881 CTGGTCCCATCAACGCACTCATCTGCGAGAGCCATACAAATGTCCAGAAATGTGGC 2940
Db 2881 CTGGTCCCATCAACGCACTCATCTGCGAGAGCCATACAAATGTCCAGAAATGTGGC 2940
Qy 2941 AAGTCTTCTCTCGGCTGCAAAATCTCGTCCGGCACCAACGTTACTCACACCGGGGAG 3000
Db 2941 AAGTCTTCTCTCGGCTGCAAAATCTCGTCCGGCACCAACGTTACTCACACCGGGGAG 3000
Qy 3001 CCTATGCTTGTCCGGAATGTGGTAAGTCTTTCAGCGCAGCGATAACCTGCTGGCCAC 3060
Db 3001 CCTATGCTTGTCCGGAATGTGGTAAGTCTTTCAGCGCAGCGATAACCTGCTGGCCAC 3060
Qy 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAATGCCAGAGTGGCGAAATCTTTTAC 3120
Db 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAATGCCAGAGTGGCGAAATCTTTTAC 3120
Qy 3121 CAGGCGGCGCACTGGCCAGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
Db 3121 CAGGCGGCGCACTGGCCAGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180

```

QY 3181 CCAGAAATGGCAAGTCTTCTCTCGGCTGCAAAATCTCGTCCGGCACCACAGTACTCAC 3240
Db |||||
QY 3181 CCAGAAATGGCAAGTCTTCTCTCGGCTGCAAAATCTCGTCCGGCACCACAGTACTCAC 3240
Db |||||
QY 3241 ACCGGTAAAAAATAGTGGCCAGCGCCGCGAGTACCGGTACGAGTTCCGGACTACGCT 3300
Db |||||
QY 3241 ACCGGTAAAAAATAGTGGCCAGCGCCGCGAGTACCGGTACGAGTTCCGGACTACGCT 3300
Db |||||

RESULT 2
AAD11589
ID AAD11589 standard; DNA; 3300 BP.
XX
AC AAD11589;
XX
DT 24-SEP-2001 (first entry)
XX
DE Partial sequence of pMal-m2 and ZFPm2 DNA.
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
KW modulation; plant technology; agriculture; ds.
XX
OS Unidentified.
XX
PH Key
FT CDS
FT /tag= a
FT /product= "ZFPm2 protein"
FT /note= "CDS does not include start and stop codon"
FT /partial
FT primer_bind
FT 2740..2790
FT /tag= b
FT /bound_moiety= "F1-f2 primer"
FT 2770..2850
FT /tag= c
FT /bound_moiety= "F1-f1 primer"
FT 2824..2889
FT /tag= d
FT primer_bind
FT 2867..2940
FT /tag= e
FT /bound_moiety= "F2-b primer"
FT 2916..2973
FT /tag= f
FT primer_bind
FT 2953..3021
FT /tag= g
FT /bound_moiety= "F3-b2 primer"
FT 2992..3042
FT /tag= h
FT primer_bind
FT 3022..3102
FT /tag= i
FT primer_bind
FT 3076..3141
FT /tag= j
FT primer_bind
FT 3119..3192
FT /tag= k
FT primer_bind
FT 3168..3225
FT /tag= l
FT primer_bind
FT 3205..3273
FT /tag= m
FT /bound_moiety= "F6-b2 primer"

WO200152620-A2.
26-JUL-2001.
19-JAN-2001; 2001WO-US001817.
XX

```

```

PR 21-JAN-2000; 2000US-0177468P.
PR 21-JUL-2000; 2000US-00620897.
XX
PA (SCRI ) SCRIPPS RES INST.
PA (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
PI Barbas CF, Stege JT, Guan X, Dalmia B;
XX
DR WPI; 2001-465325/50.
DR P-FSDB; AAE06002.
XX
PT New zinc finger proteins, useful for modulating or regulating gene
PT expression and metabolic pathways in plants, e.g. for treating in the
PT plant cells a disorder that is associated with abnormal expression of the
PT target gene.
XX
PS Example 4; Page 140-142; 156pp; English.
XX
CC The patent discloses methods and compositions to modulate the expression
CC of a target gene in plant cells. The method involves providing plant
CC cells with a zinc finger protein (ZFP) which is capable of specifically
CC binding to a target nucleotide sequence or its complementary strand
CC within a target gene and allowing the ZFP binding to the target
CC nucleotide sequence, where the expression of the target gene in the plant
CC cells is modulated. The ZFP and fusions of the ZFP proteins are useful
CC for modulating or regulating gene expression and metabolic pathways in
CC plants. The ZFP, fusion proteins and methods are useful in plant and
CC agricultural technology. The method is useful particularly for treating a
CC disorder in the plant cells, where the disorder is associated with
CC abnormal expression of the target gene. The present DNA sequence is the
CC partial sequence of pMal-m2 and ZFPm2 DNA
XX
SQ Sequence 3300 BP; 847 A; 874 C; 876 G; 703 T; 0 U; 0 Other;
Query Match 100.0%; Score 3300; DB 4; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCGACACATCGAATGGTGCAAAACCTTTCGGGTATGGCATGATAGCGCCCGAGAGA 60
Db |||||
QY 1 CCGACACATCGAATGGTGCAAAACCTTTCGGGTATGGCATGATAGCGCCCGAGAGA 60
Db |||||
QY 61 GTCAAATTCAGGGTGGTGAATGTGAACACAGTATACGATGTGCGAGAGTATGCCG 120
Db |||||
QY 61 GTCAAATTCAGGGTGGTGAATGTGAACACAGTATACGATGTGCGAGAGTATGCCG 120
Db |||||
QY 121 GTGTCTCTTATCAGACCGTTTCCGGGTGTGAACACAGTATACGATGTGCGAGAGTATGCCG 180
Db |||||
QY 121 GTGTCTCTTATCAGACCGTTTCCGGGTGTGAACACAGTATACGATGTGCGAGAGTATGCCG 180
Db |||||
QY 181 CGCGGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTAATTCACACCGGTGGCAC 240
Db |||||
QY 181 CGCGGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTAATTCACACCGGTGGCAC 240
Db |||||
QY 241 AACAACTGGCGGGCAAAACAGTCTGTGATGGCGTGGCCACTCCAGTCTGGCCCTGC 300
Db |||||
QY 241 AACAACTGGCGGGCAAAACAGTCTGTGATGGCGTGGCCACTCCAGTCTGGCCCTGC 300
Db |||||
QY 301 ACGCGCGTGCAGAAATGTTCGGCGGATTAATTAATTCGGCGCGATCACTGGGTGCCAGCG 360
Db |||||
QY 301 ACGCGCGTGCAGAAATGTTCGGCGGATTAATTAATTCGGCGCGATCACTGGGTGCCAGCG 360
Db |||||
QY 361 TGGTGGTGTGATGGTAGAACGAAAGCGCGTCAAGCCTGTAAAGCGCGGTGCACAATC 420
Db |||||
QY 421 TTCTCGGGCAACCGGTGAGTGGCTGATCAATTAATTCGGTGGATGACCAAGATGCCA 480
Db |||||
QY 421 TTCTCGGGCAACCGGTGAGTGGCTGATCAATTAATTCGGTGGATGACCAAGATGCCA 480
Db |||||
QY 481 TTGCTGTGGAAGCTGCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCCTGACAGA 540
Db |||||
QY 481 TTGCTGTGGAAGCTGCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCCTGACAGA 540
Db |||||

```

QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGGCATGGGCGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGGCATGGGCGTGGAGCATC 600
QY 601 TGGTCGCATTGGGTCACAGCAAAATCGCGCTGTAGCGGCCCATTAAGTTCTGTCGG 660
DB 601 TGGTCGCATTGGGTCACAGCAAAATCGCGCTGTAGCGGCCCATTAAGTTCTGTCGG 660
QY 661 CGCGTCTGGCTGGCTGGCTGGCTAAATATCTCACTCGCAATCAAAATTCAGCGATAG 720
DB 661 CGCGTCTGGCTGGCTGGCTGGCTAAATATCTCACTCGCAATCAAAATTCAGCGATAG 720
QY 721 CGGAACGGGAAGCGCATGTGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
DB 721 CGGAACGGGAAGCGCATGTGAGTGCCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTCCCACTGGATGCTGGTTGCCAAGCATCAGATGGCGTGGCGCAA 840
DB 781 ATGAGGGCATCGTCCCACTGGATGCTGGTTGCCAAGCATCAGATGGCGTGGCGCAA 840
QY 841 TCGCGGCCATTACCGAGTCCGGCTCGCGTTGGTGGGATATCTCGGTAGTGGGATCG 900
DB 841 TCGCGGCCATTACCGAGTCCGGCTCGCGTTGGTGGGATATCTCGGTAGTGGGATCG 900
QY 901 ACGATACCGAAGACAGCTCATGTATATCCCGCGTTAAACCAATCAAAAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTATATCCCGCGTTAAACCAATCAAAAGGATTTTC 960
QY 961 GCCTGCTGGGGCAACAGCGTGGACCGCTTGCCTGCAACTCTCAGGCCCAGGCGGTGA 1020
DB 961 GCCTGCTGGGGCAACAGCGTGGACCGCTTGCCTGCAACTCTCAGGCCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGCTGAAAGAAACCAACCTGGCGCCCAATA 1080
DB 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGCTGAAAGAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGATTCAATTAATGCACTGGCAACAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGATTCAATTAATGCACTGGCAACAGGTTT 1140
QY 1141 CCGGACTGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCTCATTTAG 1200
DB 1141 CCGGACTGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCTCATTTAG 1200
QY 1201 GCACAATCTCATGTTTGACAGCTTATCATCGACTGCAACGCTTCTCTGGCG 1260
DB 1201 GCACAATCTCATGTTTGACAGCTTATCATCGACTGCAACGCTTCTCTGGCG 1260
QY 1261 TCAGGCAGGCATCGGAAGCTGTGGTATGGCTGTGCAAGTCTGTAATCACTGCATATTCG 1320
DB 1261 TCAGGCAGGCATCGGAAGCTGTGGTATGGCTGTGCAAGTCTGTAATCACTGCATATTCG 1320
QY 1321 TGTGCTCAAGGCGCACTCCGTTCTGGATATGTTTTTTTGGCGCAATCATACGTT 1380
DB 1321 TGTGCTCAAGGCGCACTCCGTTCTGGATATGTTTTTTTGGCGCAATCATACGTT 1380
QY 1381 CTGGCAATATTTCTGAAATGAGCTGTGACAAATTAATCATCGCTCGTATATGTTGGA 1440
DB 1381 CTGGCAATATTTCTGAAATGAGCTGTGACAAATTAATCATCGCTCGTATATGTTGGA 1440
QY 1441 ATTGTGAGCGGATAACAAATTTACAGGAACAGGCGAGTCCGTTTGGTGTTCACGA 1500
DB 1441 ATTGTGAGCGGATAACAAATTTACAGGAACAGGCGAGTCCGTTTGGTGTTCACGA 1500
QY 1501 GCACTTCAACCAACAGGACCATAGATTATGAAATCTGAAGAGGTAACTGGTAATCTGG 1560
DB 1501 GCACTTCAACCAACAGGACCATAGATTATGAAATCTGAAGAGGTAACTGGTAATCTGG 1560
QY 1561 ATTACGGGCAATTAAGGCTATTAACCGTCTCGCTGAAATCGGTAAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTACGGGCAATTAAGGCTATTAACCGTCTCGCTGAAATCGGTAAAGAAATTCGAGAAAGAT 1620

QY 1621 ACCGGAATTAAGTCAACGTTGAGCATCCGATAAACTGGGAAGAAATTTCCACAGTT 1680
DB 1621 ACCGGAATTAAGTCAACGTTGAGCATCCGATAAACTGGGAAGAAATTTCCACAGTT 1680
QY 1681 CGCGCAACTGCGGATGGCCCTGACATTTCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
DB 1681 CGCGCAACTGCGGATGGCCCTGACATTTCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCGGACAAGCGTTTCCAGGACAAGCTGTAT 1800
DB 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCGGACAAGCGTTTCCAGGACAAGCTGTAT 1800
QY 1801 CGGTTTACCTGGGATGGCTAGCTTACAAACGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
DB 1801 CGGTTTACCTGGGATGGCTAGCTTACAAACGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
QY 1861 GAAGGTTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTTGGAA 1920
DB 1861 GAAGGTTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTTGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGTAAGAGCGCTGATGTTCAAC 1980
DB 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGTAAGAGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAAACGCTACTTCACTGGCCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
DB 1981 CTGCAAGAAACGCTACTTCACTGGCCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAACGCAAGTACGACATTAAGACGTTGGCGTGGATAACGCTGGCGCGAAAGCG 2100
DB 2041 TATGAAAACGCAAGTACGACATTAAGACGTTGGCGTGGATAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAACCAACACATGAAATGCAGACACCGATTAC 2160
DB 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAACCAACACATGAAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAAAGCTGCGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCGCTGG 2220
DB 2161 TCCATCGCAGAAAGCTGCGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCGCTGG 2220
QY 2221 GCATGGTCCACATCGACACCGCAAGTGAATTAATGTTAAACGCTACTGCGACCTTC 2280
DB 2221 GCATGGTCCACATCGACACCGCAAGTGAATTAATGTTAAACGCTACTGCGACCTTC 2280
QY 2281 AAGGCTCAACCATCAAAACCGTTTCTGCGCTGCTGAGCGCAGGATTAACCGCCGCTAG 2340
DB 2281 AAGGCTCAACCATCAAAACCGTTTCTGCGCTGCTGAGCGCAGGATTAACCGCCGCTAG 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAATACTATCTGCTGACTGATGAAGTCTG 2400
DB 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAATACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GNAAGGTTTATAAGACAAACCGCTGGTGGCTAGCGCTGAAGTCTTACCAGGAAGAG 2460
DB 2401 GNAAGGTTTATAAGACAAACCGCTGGTGGCTAGCGCTGAAGTCTTACCAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCACGTATTGCCGCCACCATGGAATAACCGCCAGAAAGGTGAAATCATG 2520
DB 2461 TTGGCGAAAGATCCACGTATTGCCGCCACCATGGAATAACCGCCAGAAAGGTGAAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATCGCGTACTGCGGTGATCAACGCGC 2580
DB 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATCGCGTACTGCGGTGATCAACGCGC 2580
QY 2581 GCCAGGCTGCTCAGACTGTGATGAAGCCCTGAAAGACCGCGCAGACTTAATTCGAGCTCG 2640
DB 2581 GCCAGGCTGCTCAGACTGTGATGAAGCCCTGAAAGACCGCGCAGACTTAATTCGAGCTCG 2640
QY 2641 AACACAAACAAATAACAAACCTCGGGATCGAGGAGGAGGATTTTCAGAAATTC 2700
DB 2641 AACACAAACAAATAACAAACCTCGGGATCGAGGAGGAGGATTTTCAGAAATTC 2700
QY 2701 GGATCCTCTTCTGCTGGCCCGCGGCTCGAGCCCGGGGAGAAAGCCCTATGCTTGT 2760

```
Db 2701 GGATCCTCTCTCTGTGGCCAGCGGCCCTCGAGCCGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTGGTAAGTCTCTCTCAGAGCTCTACCTGGTGGCCACACCGGTACCCAC 2820
Db 2761 CCGGAATGTGGTAAGTCTCTCTCAGAGCTCTACCTGGTGGCCACACCGGTACCCAC 2820
QY 2821 ACGGGTGAACAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGGTGAACAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGGTGGCCATCAACGCACTCATACTGGGAGAGCCATACAAATGTCGAGATGTGGC 2940
Db 2881 CTGGTGGCCATCAACGCACTCATACTGGGAGAGCCATACAAATGTCGAGATGTGGC 2940
QY 2941 AAGTCTTTCTCTCGGTCTGACAATCTCGTCCGGCACCACGCTACTCACACGGGGAGAG 3000
Db 2941 AAGTCTTTCTCTCGGTCTGACAATCTCGTCCGGCACCACGCTACTCACACGGGGAGAG 3000
QY 3001 CCCTATGTTGTCGGAAATGGTAAAGTCTTTTCAGCCGACGAGTAACTGCTGGCGCCAC 3060
Db 3001 CCCTATGTTGTCGGAAATGGTAAAGTCTTTTCAGCCGACGAGTAACTGCTGGCGCCAC 3060
QY 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
QY 3121 CAGGCGCGCCACTGGCCAGCCATCAACGCACTCATCTGGGAGAGCCATACAAATGT 3180
Db 3121 CAGGCGCGCCACTGGCCAGCCATCAACGCACTCATCTGGGAGAGCCATACAAATGT 3180
QY 3181 CCAGAATGCGCAAGTCTTTCTCTCGGTCTGACAATCTCGTCCGGCACCACGCTACTCAC 3240
Db 3181 CCAGAATGCGCAAGTCTTTCTCTCGGTCTGACAATCTCGTCCGGCACCACGCTACTCAC 3240
QY 3241 ACCGGTAAAAAAGTGGCCAGCCGCGGACGCTAGTACCGGTACGAGTTCGGGACTACGCT 3300
Db 3241 ACCGGTAAAAAAGTGGCCAGCCGCGGACGCTAGTACCGGTACGAGTTCGGGACTACGCT 3300

RESULT 3
ID AAD11592 standard; DNA; 3300 BP.
XX
AC AAD11592;
XX
DT 24-SEP-2001 (first entry)
XX
DE Partial sequence of pMal-Ap3 and ZFPap3 DNA.
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
KW modulation; plant technology; agriculture; Ap3; APETALA3; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 2719..3270
FT /tag= a
FT /product= "ZFPap3 protein"
FT /transl_except= (pos:3124..3126, aa:Ser)
FT /transl_except= (pos:3127..3129, aa:Ser)
FT /transl_except= (pos:3136..3138, aa:Val)
FT /note= "CDS does not include start and stop codon"
FT /partial
FT primer_bind 2740..2790
FT /tag= b
FT /bound_moietly= "F1-f2 primer"
FT primer_bind 2770..2850
FT /tag= c
FT /bound_moietly= "F1-f1 primer"
FT primer_bind 2824..2889
FT /tag= d
FT /bound_moietly= "F2-b primer"
FT
```

```
FT primer_bind 2867..2940
FT /tag= e
FT /bound_moietly= "F2-f primer"
FT primer_bind 2916..2973
FT /tag= f
FT /bound_moietly= "F3-b1 primer"
FT primer_bind 2953..3021
FT /tag= g
FT /bound_moietly= "F3-b2 primer"
FT primer_bind 2992..3042
FT /tag= h
FT /bound_moietly= "F4-f2 primer"
FT primer_bind 3022..3102
FT /tag= i
FT /bound_moietly= "F4-f1 primer"
FT primer_bind 3076..3141
FT /tag= j
FT /bound_moietly= "F5-b primer"
FT primer_bind 3119..3192
FT /tag= k
FT /bound_moietly= "F5-f primer"
FT primer_bind 3169..3225
FT /tag= l
FT /bound_moietly= "F6-b1 primer"
FT primer_bind 3205..3273
FT /tag= m
FT /bound_moietly= "F6-b2 primer"
XX
XX WO200152620-A2.
XX
XX 26-JUL-2001.
XX
XX 19-JAN-2001; 2001WO-US001817.
XX
XX 21-JAN-2000; 2000US-0177488P.
XX
XX 21-JUL-2000; 2000US-00620897.
XX
XX (SCRI ) SCRIPPS RES INST.
XX
XX (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
XX Barbas CP, Stege JT, Guan X, Dalmia B;
XX
XX WPI; 2001-465325/50.
XX
XX P-PSDB; AAE06005.
XX
XX New zinc finger proteins, useful for modulating or regulating gene
XX expression and metabolic pathways in plants, e.g. for treating in the
XX plant cells a disorder that is associated with abnormal expression of the
XX target gene.
XX
XX Example 4; Page 148-149; 156pp; English.
XX
XX The patent discloses methods and compositions to modulate the expression
XX of a target gene in plant cells. The method involves providing plant
XX cells with a zinc finger protein (ZFP) which is capable of specifically
XX binding to a target nucleotide sequence or its complementary strand
XX within a target gene and allowing the ZFP binding to the target
XX nucleotide sequence, where the expression of the target gene in the plant
XX cells is modulated. The ZFP and fusions of the ZFP proteins are useful
XX for modulating or regulating gene expression and metabolic pathways in
XX plants. The ZFP, fusion proteins and methods are useful in plant and
XX agricultural technology. The method is useful particularly for treating a
XX disorder in the plant cells, where the disorder is associated with
XX abnormal expression of the target gene. The present DNA sequence is the
XX partial sequence of pMal-Ap3 (APETALA3) and ZFPap3 DNA
XX
XX Sequence 3300 BP; 853 A; 872 C; 876 G; 699 T; 0 U; 0 Other;
SQ
Query Match 97.8%; Score 3226.4; DB 4; Length 3300;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3254; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGTTGCAAAACCTTCGCGGTATGGCATATAGGCCCGGAGAGA 60
```

Db 1 CCGACACCATCGAATGGTGCAGAACTTTTCGGGATATGGCATGATAGCGCCCGGAAGAGA 60
Qy 61 GTCAATTCAAGGTGTGAATGTGAACACAGTAAACGTTATACGATGTCGACAGTATGCCG 120
Db 61 GTCAATTCAAGGTGTGAATGTGAACACAGTAAACGTTATACGATGTCGACAGTATGCCG 120
Qy 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACGCGCCAGCCAGCTTTCTCGAANA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACGCGCCAGCCAGCTTTCTCGAANA 180
Qy 181 CCGGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTAACATTCCCAACCGCTGGGCAC 240
Db 181 CCGGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTAACATTCCCAACCGCTGGGCAC 240
Qy 241 AACAACTGGCGGCAAAACAGTCTGTGTGCTGATGGCGTTGCCACTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAAAACAGTCTGTGTGCTGATGGCGTTGCCACTCCAGTCTGGCCCTGC 300
Qy 301 AGCGCCGTGCGCAAAATGTTCGGGCGATTAATCTCGGCGCGATCAACTGGGTGCCAGG 360
Db 301 AGCGCCGTGCGCAAAATGTTCGGGCGATTAATCTCGGCGCGATCAACTGGGTGCCAGG 360
Qy 361 TGGTGTGTTCGATGTAGAACGAAGCGCGTGCAGAGCTGTAAAGCGGCGTGCACAATC 420
Db 361 TGGTGTGTTCGATGTAGAACGAAGCGGCGTGCAGAGCTGTAAAGCGGCGTGCACAATC 420
Qy 421 TTCTCGCGCAACGCTCAGTGGGTGATCATTAACATCTCCGTGGATGACACGAGTGCCA 480
Db 421 TTCTCGCGCAACGCTCAGTGGGTGATCATTAACATCTCCGTGGATGACACGAGTGCCA 480
Qy 481 TTGCTGTGGAGCTGCCTGCACATAATGTTCCGGCGTTATTTCTTGATGCTCTGACACGA 540
Db 481 TTGCTGTGGAGCTGCCTGCACATAATGTTCCGGCGTTATTTCTTGATGCTCTGACACGA 540
Qy 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGTACGGACTGGCGGTGAAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGTACGGACTGGCGGTGAAGCATC 600
Qy 601 TGGTGCATTTGGTCCAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCTGTCTCGG 660
Db 601 TGGTGCATTTGGTCCAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCTGTCTCGG 660
Qy 661 CGGCTCGGCTGCGTGGCGATTAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db 661 CGGCTCGGCTGCGTGGCGATTAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Qy 721 CCGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CCGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
Db 781 ATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
Qy 841 TCGCGGCAATTAACGAGTCCGGGCTGCGGTTGGTGGCGGATATCTCGGTAGTGGGATAG 900
Db 841 TCGCGGCAATTAACGAGTCCGGGCTGCGGTTGGTGGCGGATATCTCGGTAGTGGGATAG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAAAAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAAAAGGATTTTC 960
Qy 961 GCCTGCTGGGCAAAACAGCGTGAACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAAAACAGCGTGAACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGTCGGCTCTCACTGGTGAAGAAAAACCACTTCGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGTCGGCTCTCACTGGTGAAGAAAAACCACTTCGGCGCCCAATA 1080
Qy 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCTGGCAGCAGAGTTT 1140
Db 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCTGGCAGCAGAGTTT 1140

Db 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCTGGCAGCAGAGTTT 1140
Qy 1141 CCGCATGGAAGCGGCGAGTGAAGCGCAACCCAAATTAATGTAGTTAGCTCACTCATTTAG 1200
Db 1141 CCGCATGGAAGCGGCGAGTGAAGCGCAACCCAAATTAATGTAGTTAGCTCACTCATTTAG 1200
Qy 1201 GCACAATTCATGTTTGACAGCTTATCATCGATGCAAGCGTGCACCAATGCTTCGGCG 1260
Db 1201 GCACAATTCATGTTTGACAGCTTATCATCGATGCAAGCGTGCACCAATGCTTCGGCG 1260
Qy 1261 TCAGGACGCCATTCGGAAGCTGTGGTGTGAGGTGCTGAGGTGCTTAATCACTGCATAAATTCG 1320
Db 1261 TCAGGACGCCATTCGGAAGCTGTGGTGTGAGGTGCTGAGGTGCTTAATCACTGCATAAATTCG 1320
Qy 1321 TGTCTCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTTCGCGCAATCAATAACGTT 1380
Db 1321 TGTCTCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTTCGCGCAATCAATAACGTT 1380
Qy 1381 CTGGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
Db 1381 CTGGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
Qy 1441 ATTGTAGCGGATAACAAATTTTCAACAGCAACACGCAAGTCCGTTTTCAGTATCTCG 1500
Db 1441 ATTGTAGCGGATAACAAATTTTCAACAGCAACACGCAAGTCCGTTTTCAGTATCTCG 1500
Qy 1501 GCATTTCACCAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTATATCTCG 1560
Db 1501 GCATTTCACCAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTATATCTCG 1560
Qy 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAGTTCGGTGAAGATTCGAGAAAGT 1620
Db 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAGTTCGGTGAAGATTCGAGAAAGT 1620
Qy 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGATATACTGGAAGAGAAATTCGCCAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGATATACTGGAAGAGAAATTCGCCAGGTT 1680
Qy 1681 CGGCAACTGCGATGCGCTTGACATTAATCTTCTGGGCAACGACCCGTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGCGATGCGCTTGACATTAATCTTCTGGGCAACGACCCGTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCGGCAAGAGGTTCCAGACAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCGGCAAGAGGTTCCAGACAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAAGGCAAGCTGATTGCTTACCGCATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAAGGCAAGCTGATTGCTTACCGCATCGCTGT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCGGCGCTGGATTAAGAACTGAAAGCAAGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGGCGCTGGATTAAGAACTGAAAGCAAGTAAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCCACCTGCGCTGATGCTGCTGACGGGGTTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCCACCTGCGCTGATGCTGCTGACGGGGTTTATGCGTTCAAG 2040
Qy 2041 TATGAAAAACGCAAGTACGACATTAAGAACGCTGGGCGTGGATTAACGCTGGGCGGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGAACGCTGGGCGTGGATTAACGCTGGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACATGAAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACAACATGAAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGCAACAGCGATGACCATCAAGGCGCCGTTG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGCAACAGCGATGACCATCAAGGCGCCGTTG 2220


```
QY 2221 GCATGGTCCAAATCGACACACGAAAGTGAATTATGGTAAACGGTACTGCGGACCTTC 2280
Db |||||
QY 2221 GCATGGTCCAAATCGACACACGAAAGTGAATTATGGTAAACGGTACTGCGGACCTTC 2280
Db |||||
QY 2281 AAGGTCACACCATCAAAACCGTTCGTTGGCTGCTGAGCGGACGATTAACCGCCCACT 2340
Db |||||
QY 2281 AAGGTCACACCATCAAAACCGTTCGTTGGCTGCTGAGCGGACGATTAACCGCCCACT 2340
Db |||||
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACGTGAGGTCCTG 2400
Db |||||
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACGTGAGGTCCTG 2400
Db |||||
QY 2401 GAAGCGGTAAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db |||||
QY 2401 GAAGCGGTAAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db |||||
QY 2461 TTGGCGAAGATCCACGATTTGCGCGCCACCATGGAAGACCGCCGAGAAAGTGAATCATG 2520
Db |||||
QY 2461 TTGGCGAAGATCCACGATTTGCGCGCCACCATGGAAGACCGCCGAGAAAGTGAATCATG 2520
Db |||||
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTGATGCGGTGCTGCTGCGGTGATCAACGCC 2580
Db |||||
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTGATGCGGTGCTGCTGCGGTGATCAACGCC 2580
Db |||||
QY 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db |||||
QY 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db |||||
QY 2641 AACACACACACATTAACAATAACAACCTCGGATCGAGGAGGAGATTCAGAAATTC 2700
Db |||||
QY 2641 AACACACACACATTAACAATAACAACCTCGGATCGAGGAGGAGATTCAGAAATTC 2700
Db |||||
QY 2701 GGATCCTCTTCTGTCGCGCCAGCGGCTCTGAGCCCGGGAGAGCCCTATGCTGTG 2760
Db |||||
QY 2701 GGATCCTCTTCTGTCGCGCCAGCGGCTCTGAGCCCGGGAGAGCCCTATGCTGTG 2760
Db |||||
QY 2761 CCGGAATGTGTAAAGTCTTCTCAGAGCTCTCAGAGTCTGCGCCACAGCGTACCCAC 2820
Db |||||
QY 2761 CCGGAATGTGTAAAGTCTTCTCAGAGCTCTCAGAGTCTGCGCCACAGCGTACCCAC 2820
Db |||||
QY 2821 ACAGGTGAAACACCGTAAATGCGCAGAGTGGCGCAATCTTTAGCAGTCCAGCAAC 2880
Db |||||
QY 2821 ACAGGTGAAACACCGTAAATGCGCAGAGTGGCGCAATCTTTAGCAGTCCAGCAAC 2880
Db |||||
QY 2881 CTGGTGGCCATCAACGACCTCATCTGCGGAGAGCCATACAAATGTCCAGAAATGTGGC 2940
Db |||||
QY 2881 CTGGTGGCCATCAACGACCTCATCTGCGGAGAGCCATACAAATGTCCAGAAATGTGGC 2940
Db |||||
QY 2941 AAGTCTTTCTCGGTCTGACAACTCTGTCGCGCACCAAGTACTCACCGGGGAGAAG 3000
Db |||||
QY 2941 AAGTCTTTCTCGGTCTGACAACTCTGTCGCGCACCAAGTACTCACCGGGGAGAAG 3000
Db |||||
QY 3001 CCCTATGCTTCTCGGATGTTAGTCTTCTGAGCGGAGTACCTGCTGCGGAC 3060
Db |||||
QY 3001 CCCTATGCTTCTCGGATGTTAGTCTTCTGAGCGGAGTACCTGCTGCGGAC 3060
Db |||||
QY 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCGGAGTGGCGCAATCTTTAGC 3120
Db |||||
QY 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCGGAGTGGCGCAATCTTTAGC 3120
Db |||||
QY 3121 CAGCGCGCCACCTGGCAGGCATCAACGCACTCATCTGCGGAGAGGCAATCAAAATGT 3180
Db |||||
QY 3121 CAGCGCGCCACCTGGCAGGCATCAACGCACTCATCTGCGGAGAGGCAATCAAAATGT 3180
Db |||||
QY 3181 CCAGAAATGTGGCAAGTCTTTCTCTCGGTCTGACAACTCTGTCGCGCACCAACGCTACTCAC 3240
Db |||||
QY 3181 CCAGAAATGTGGCAAGTCTTTCTCTCGGTCTGACAACTCTGTCGCGCACCAACGCTACTCAC 3240
Db |||||
QY 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGGCGGAGTACCGGTACGACGTTCCGGACTAGCT 3300
Db |||||
QY 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGGCGGAGTACCGGTACGACGTTCCGGACTAGCT 3300
Db |||||
```

```
RESULT 4
AAD11591
ID AAD11591 standard; DNA; 3300 BP.
XX
AC AAD11591;
XX
DT 24-SEP-2001 (first entry)
XX
DE Partial sequence of pMal-m4 and ZFPm4 DNA.
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
modulation; plant technology; agriculture; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 2719..3270
FT FT /*tag= a
FT FT /product= "ZFPm4 protein"
FT FT /transl_except= (pos:3046..3048, aa:Ser)
FT FT /note= "CDS does not include start and stop codon"
FT FT /partial
FT primer_bind 2740..2790
FT FT /*tag= b
FT FT /bound_moiety= "F1-f2 primer"
FT FT 2770..2850
FT FT /*tag= c
FT FT /bound_moiety= "F1-f1 primer"
FT FT 2824..2889
FT FT /*tag= d
FT FT /bound_moiety= "F2-b primer"
FT FT 2867..2940
FT FT /*tag= e
FT FT /bound_moiety= "F2-f primer"
FT FT 2916..2973
FT FT /*tag= f
FT FT /bound_moiety= "F3-b1 primer"
FT FT 2953..3021
FT FT /*tag= g
FT FT /bound_moiety= "F3-b2 primer"
FT FT 2992..3042
FT FT /*tag= h
FT FT /bound_moiety= "F4-f2 primer"
FT FT 3022..3102
FT FT /*tag= i
FT FT /bound_moiety= "F4-f1 primer"
FT FT 3076..3141
FT FT /*tag= j
FT FT /bound_moiety= "F5-b primer"
FT FT 3119..3192
FT FT /*tag= k
FT FT /bound_moiety= "F5-f primer"
FT FT 3168..3225
FT FT /*tag= l
FT FT /bound_moiety= "F6-b1 primer"
FT FT 3205..3273
FT FT /*tag= m
FT FT /bound_moiety= "F6-b2 primer"
XX
PN WO200152620-A2.
XX
PD 26-JUL-2001.
XX
PF 19-JAN-2001; 2001WO-US001817.
XX
PR 21-JAN-2000; 2000US-0177468P.
XX
PR 21-JUL-2000; 2000US-0062089P.
XX
PA (SCRI ) SCRIPPS RES INST.
XX
PA (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
PI Barbas CF, Stege JT, Guan X, Dalmia B;
```



```
QY 1741 GCTCAATCTGGCTGTGGCTGAATTCACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAATTCACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCGCTGACGTTACAAACGGAAGCTGATGCTTACCCGATCCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGCTGACGTTACAAACGGAAGCTGATGCTTACCCGATCCTGTT 1860
QY 1861 GAAGCGTTATCCTGATTTATACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCCTGATTTATACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGGTGATGCTGTGACGGGGTTATGCTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGGTGATGCTGTGACGGGGTTATGCTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGGTGATGATGCGGCGGAAACGG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGGTGATGATGCGGCGGAAACGG 2100
QY 2101 GGTCTGACCTTCTGCTGATTAAGAAACCAACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGATTAAGAAACCAACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCGCGTGG 2220
Db 2161 TCCATCGCAAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCGCGTGG 2220
QY 2221 GCATGCTCAACATCGACACAGCAAGTGAATTAATGCTGTAACCGTACCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACAGCAAGTGAATTAATGCTGTAACCGTACCGGACCTTC 2280
QY 2281 AAGGCTCAACCATCAAAACCGTTCTGTCGCTGAGCGAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGCTCAACCATCAAAACCGTTCTGTCGCTGAGCGAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGATGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGATGATGAAGTCTG 2400
QY 2401 GAAAGCGTTAATAAGACAAACCGCTGGCTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAAGCGTTAATAAGACAAACCGCTGGCTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGTTATTCGCCCAACCATGGAACACGCCCAAGAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATTCGCCCAACCATGGAACACGCCCAAGAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGAGATGCTCGCTTCTGCTGATGCGGTGCGTACTCGGCTGATCAAGCC 2580
Db 2521 CCGAACATCCCGAGATGCTCGCTTCTGCTGATGCGGTGCGTACTCGGCTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGTCTG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGTCTG 2640
QY 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
QY 2701 GGATCTCTTCTCTGTCGCGGCGGCGCTCGAGCGCGGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTCGCGGCGGCGCTCGAGCGCGGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTGTAAAGTCTTCTCTCAGAGCTCTCAGCTGTGCGCCACACGAGCTACCCAC 2820
Db 2761 CCGGAATGTGTAAAGTCTTCTCTCAGAGCTCTCAGCTGTGCGCCACACGAGCTACCCAC 2820
```

```
QY 2821 ACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTAGCAGTCCAGCAAC 2880
Db 2821 ACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTAGCAGTCCAGCAGC 2880
QY 2881 CTGGTGGGCCATCAACCGCACTCATCTGGCGAGAAGCCATACAAATGTCCAGAAATGTGGC 2940
Db 2881 CTGGTGGGCCATCAACCGCACTCATCTGGCGAGAAGCCATACAAATGTCCAGAAATGTGGC 2940
QY 2941 AAGTCTTTCTCTCGGTCTGACAAATCTCTCGCGGACCAACGTAATCTCACACCGGGAGAAG 3000
Db 2941 AAGTCTTTCTCTCGGTCTGACAAATCTCTCGCGGACCAACGTAATCTCACACCGGGAGAAG 3000
QY 3001 CCCTATCTTCTGTCGGGAATGTGTAAGTCTTCTAGCCGCGAGGATACCTGGTGGCCAC 3060
Db 3001 CCCTATCTTCTGTCGGGAATGTGTAAGTCTTCTAGCCGCGAGGATACCTGGTGGCCAC 3060
QY 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTAGC 3120
QY 3121 CAGCGCGGCACTGGCGGAGCCATCAACGCACTCATCTGCGGAGAAGCCATACAAATGT 3180
Db 3121 CAGCGCGGCACTGGCGGAGCCATCAACGCACTCATCTGCGGAGAAGCCATACAAATGT 3180
QY 3181 CCAGAAATGTGGCAAGTCTTCTCTCGGTCTGACAAATCTGTCGGGCAACGTAATCTTAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTCTCTCGGTCTGACAAATCTGTCGGGCAACGTAATCTTAC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCGGCGGCGGAGTACCGGTACGACGTTCCGGGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCGGCGGCGGAGTACCGGTACGACGTTCCGGGACTACGCT 3300

RESULT 5
AAD11590
ID AAD11590 standard; DNA; 3300 BP.
XX AC AAD11590;
XX DT 24-SEP-2001 (first entry)
XX DE Partial sequence of pMal-m3 and ZFPm3 DNA.
XX KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
XX modulation; plant technology; agriculture; ds.
XX OS Unidentified.
XX FH Key Location/Qualifiers
FT CDS 2719..3270
FT FT /*tag= a
FT FT /product= "ZFPm3 protein"
FT FT /note= "CDS does not include start and stop codon"
FT FT primer_bind
FT FT 2740..2790
FT FT /*tag= b
FT FT /bound_moiety= "F1-f2 primer"
FT FT primer_bind
FT FT 2770..2850
FT FT /*tag= c
FT FT /bound_moiety= "F1-f1 primer"
FT FT primer_bind
FT FT 2824..2889
FT FT /*tag= d
FT FT /bound_moiety= "F2-b primer"
FT FT primer_bind
FT FT 2867..2940
FT FT /*tag= e
FT FT /bound_moiety= "F2-f primer"
FT FT primer_bind
FT FT 2916..2973
FT FT /*tag= f
FT FT /bound_moiety= "F3-b1 primer"
FT FT primer_bind
FT FT 2953..3021
FT FT /*tag= g
FT FT /bound_moiety= "F3-b2 primer"
FT FT primer_bind
FT FT 2992..3042
```

```
FT      /*tag= h
FT      /bound_moiety= "F4-f2 primer"
FT      3022..3102
FT      /*tag= i
FT      /bound_moiety= "F4-f1 primer"
FT      3076..3141
FT      /*tag= j
FT      /bound_moiety= "F5-b primer"
FT      3119..3192
FT      /*tag= k
FT      /bound_moiety= "F5-f primer"
FT      3168..3225
FT      /*tag= l
FT      /bound_moiety= "F6-b1 primer"
FT      3205..3273
FT      /*tag= m
FT      /bound_moiety= "F6-b2 primer"
XX
PN      WO200152620-A2.
XX
XX      26-JUL-2001.
XX
XX      19-JAN-2001; 2001WO-US001817.
XX
XX      21-JAN-2000; 2000US-0177468P.
XX      21-JUL-2000; 2000US-00620897.
XX
PA      (SCRI ) SCRIPPS RES INST.
PA      (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
XX      Barbas CF, Stege JT, Guan X, Dalmia B;
XX      WPI; 2001-465325/50.
XX      P-PSDB; AA06003.
XX
XX      New zinc finger proteins, useful for modulating or regulating gene
XX      expression and metabolic pathways in plants, e.g. for treating in the
XX      plant cells a disorder that is associated with abnormal expression of the
XX      target gene.
XX
PS      Example 4; Page 143-145; 156pp; English.
XX
XX      The patent discloses methods and compositions to modulate the expression
XX      of a target gene in plant cells. The method involves providing plant
XX      cells with a zinc finger protein (ZFP) which is capable of specifically
XX      binding to a target nucleotide sequence or its complementary strand
XX      within a target gene and allowing the ZFP binding to the target
XX      nucleotide sequence, where the expression of the target gene in the plant
XX      cells is modulated. The ZFP and fusions of the ZFP proteins are useful
XX      for modulating or regulating gene expression and metabolic pathways in
XX      plants. The ZFP, fusion proteins and methods are useful in plant and
XX      agricultural technology. The method is useful particularly for treating a
XX      disorder in the plant cells, where the disorder is associated with
XX      abnormal expression of the target gene. The present DNA sequence is the
XX      Partial sequence of pMal-m3 and ZFPm3 DNA
XX
SQ      Sequence 3300 BP; 845 A; 878 C; 877 G; 700 T; 0 U; 0 Other;
      Query Match      97.2%; Score 3207.2; DB 4; Length 3300;
      Best Local Similarity 98.2%; Score 3242; Mismatches 58; Indels 0; Gaps 0;
      Matches 3242; Conservative 0;

QY      1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGCGATGATAGGCCCGCGGAAGAGA 60
DB      1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGCGATGATAGGCCCGCGGAAGAGA 60
QY      61 GTCAATTCAGGTGGTGAATGTGAACCCAGTAAACGTTATACGATTCGCAGAGTATCGCG 120
DB      61 GTCAATTCAGGTGGTGAATGTGAACCCAGTAAACGTTATACGATTCGCAGAGTATCGCG 120
QY      121 GTGTCTCTTATCAGACCGGTTCCCGCGGTGGTGAACCCAGGCCACGCTTCTCGGAAA 180
DB      121 GTGTCTCTTATCAGACCGGTTCCCGCGGTGGTGAACCCAGGCCACGCTTCTCGGAAA 180
```

```
QY      181 CGCGGGAAAAAGTGGAAAGCGCGCGATGCGGAGCTGAATTATTCATTCCTCCAAACCGGTGGCAC 240
DB      181 CGCGGGAAAAAGTGGAAAGCGCGCGATGCGGAGCTGAATTATTCATTCCTCCAAACCGGTGGCAC 240
QY      241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300
DB      241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300
QY      301 ACGCCGCGTCGCAAAATTGTCCGCGCGATTAATATCTCCGCGCGATCAACTGGGTGCGACGC 360
DB      301 ACGCCGCGTCGCAAAATTGTCCGCGCGATTAATATCTCCGCGCGATCAACTGGGTGCGACGC 360
QY      361 TGGTGGTGTGATGCTAGAACGAAGCGCGCTGGAAGCTGTAAAGCGCGGTGCACAATC 420
DB      361 TGGTGGTGTGATGCTAGAACGAAGCGCGCTGGAAGCTGTAAAGCGCGGTGCACAATC 420
QY      421 TTCTCGGCAACCGCTCAGTGGGCTGATCACTAACTATCCGCTGATGACCAAGATGCCA 480
DB      421 TTCTCGGCAACCGCTCAGTGGGCTGATCACTAACTATCCGCTGATGACCAAGATGCCA 480
QY      481 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGCTCTCGACCGA 540
DB      481 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGCTCTCGACCGA 540
QY      541 CACCCATCAACAGTATTATTTTCTCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC 600
DB      541 CACCCATCAACAGTATTATTTTCTCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC 600
QY      601 TGGTCGCATTGGGTCCACGACCAATCGCGCTGTTAGCGGGGCCATTAAAGTTCTGTCTCGG 660
DB      601 TGGTCGCATTGGGTCCACGACCAATCGCGCTGTTAGCGGGGCCATTAAAGTTCTGTCTCGG 660
QY      661 CGCGTCTGCGCTGCGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB      661 CGCGTCTGCGCTGCGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY      721 CGGAAACGGGAAAGGCGACTGGAGTGCATGTCGGTTCGTTTCAACAAACCAATGCAATGCTGA 780
DB      721 CGGAAACGGGAAAGGCGACTGGAGTGCATGTCGGTTCGTTTCAACAAACCAATGCAATGCTGA 780
QY      781 ATGAGGCGATCGTTCACGCTCGGATGCTGGTTCGCAACGATCAGATGCGGTGGCGGCA 840
DB      781 ATGAGGCGATCGTTCACGCTCGGATGCTGGTTCGCAACGATCAGATGCGGTGGCGGCA 840
QY      841 TGGCGCCATTACCGAGTCCCGGCTGCGCGTGGTGGCGATATCTCGGTAGTGGGATACG 900
DB      841 TGGCGCCATTACCGAGTCCCGGCTGCGCGTGGTGGCGATATCTCGGTAGTGGGATACG 900
QY      901 ACGATACCGAAGACAGCTCATGTTTATATCCGCGCGTTAACCCATCAAAACAGGATTTTC 960
DB      901 ACGATACCGAAGACAGCTCATGTTTATATCCGCGCGTTAACCCATCAAAACAGGATTTTC 960
QY      961 GCCTGCTGGGCAAAACGACGCTGCAACCGCTTGGTGCACACTCTCTCAGGCCACGCGGTGA 1020
DB      961 GCCTGCTGGGCAAAACGACGCTGCAACCGCTTGGTGCACACTCTCTCAGGCCACGCGGTGA 1020
QY      1021 AGGCAATCAGCTGTTGCCCGCTCTCACTGCTGAAAGAAAACCAACCCCTGGCGCCCAATA 1080
DB      1021 AGGCAATCAGCTGTTGCCCGCTCTCACTGCTGAAAGAAAACCAACCCCTGGCGCCCAATA 1080
QY      1081 CGCAAAACCGCTCTCCCGCGCGTTGGCCGATTCATTAAATGCAGCTGGCACACAGGTTT 1140
DB      1081 CGCAAAACCGCTCTCCCGCGCGTTGGCCGATTCATTAAATGCAGCTGGCACACAGGTTT 1140
QY      1141 CCGGCTGGAAGCGGCGAGTGAGGCGAACCGCAATTAATGTAGTACTCACTCACTTAG 1200
DB      1141 CCGGCTGGAAGCGGCGAGTGAGGCGAACCGCAATTAATGTAGTACTCACTCACTTAG 1200
QY      1201 GCACAATTTCTCATGTTTGCACAGCTTATCATCATCTGCACGGTGCACCAATGCTTCTGGCG 1260
DB      1201 GCACAATTTCTCATGTTTGCACAGCTTATCATCATCTGCACGGTGCACCAATGCTTCTGGCG 1260
```

```
QY 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGCAGGTGCTAAATCACTGCATATTCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGCAGGTGCTAAATCACTGCATATTCG 1320
QY 1321 TGTGCGTCAAGGCGCACTCCGCTTCTGGATATGTTTTTGGCGGACATCAATACGGTT 1380
Db 1321 TGTGCGTCAAGGCGCACTCCGCTTCTGGATATGTTTTTGGCGGACATCAATACGGTT 1380
QY 1381 CTGGCAATATTTCTGAATAGCTGTGTGACATTAATCATCGGCTGTAATGTTGGA 1440
Db 1381 CTGGCAATATTTCTGAATAGCTGTGTGACATTAATCATCGGCTGTAATGTTGGA 1440
QY 1441 ATTGTAGCGGATAACAAATTTACACAGGAAACAGCCAGTCCGTTTAGTGTTTTACGA 1500
Db 1441 ATTGTAGCGGATAACAAATTTACACAGGAAACAGCCAGTCCGTTTAGTGTTTTACGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTAATGAAACTGAAGAGGTAAACTGGTAATCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTAATGAAACTGAAGAGGTAAACTGGTAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGAAGAGAAATTCACACAGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGAAGAGAAATTCACACAGTT 1680
QY 1681 GCGCAATCGGATGCGGCTGACATTAATCTTGGGACACAGACCGCTTTGGTGCTAC 1740
Db 1681 GCGCAATCGGATGCGGCTGACATTAATCTTGGGACACAGACCGCTTTGGTGCTAC 1740
QY 1741 GCTCAATCTGGGCTGTGGCTGAATATCAACCGGACAAAGCGTTCCAGGACAAAGCTGAT 1800
Db 1741 GCTCAATCTGGGCTGTGGCTGAATATCAACCGGACAAAGCGTTCCAGGACAAAGCTGAT 1800
QY 1801 CCGTTTACCTGGATGCGGTACGTTTAAACCGGACGAGTGTATTCACCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGGTACGTTTAAACCGGACGAGTGTATTCACCGATCGCTGTT 1860
QY 1861 GAAGGTTATCGCTGATTTATACAAAGATCTGTGCGGACCGGACCGGCAAAACCTGGAA 1920
Db 1861 GAAGGTTATCGCTGATTTATACAAAGATCTGTGCGGACCGGACCGGCAAAACCTGGAA 1920
QY 1921 GAGATCCGCGGCTGGATAAAGAACTGAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGCGGCTGGATAAAGAACTGAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACGTAATCTTCACTGCGGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACGTAATCTTCACTGCGGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGGCTGGATACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGGCTGGATACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACACATGATGACACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACACATGATGACACACCGATTAC 2160
QY 2161 TCCATCGCAAGTCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCTGG 2220
Db 2161 TCCATCGCAAGTCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCTGG 2220
QY 2221 GCATGGTCCAACTCGACACAGCAAGTGAATTAATGTTGTAACGCTGCGGACCTTC 2280
Db 2221 GCATGGTCCAACTCGACACAGCAAGTGAATTAATGTTGTAACGCTGCGGACCTTC 2280
QY 2281 AAGGTCACCAATCAACCGTTCTGGTGCTGAGCGCAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGTCACCAATCAACCGTTCTGGTGCTGAGCGCAGGTATTAACGCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGATTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
```

```
Db 2341 CCGAACAAAGAGCTGGCAAAAGATTTCTCGAAAACTATCTGCTGACTGATGAAGGCTG 2400
QY 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGTGAAGTCTTACGAGCAAGAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGTGAAGTCTTACGAGCAAGAG 2460
QY 2461 TTGGCGAAAGATTCACCGTATTTGGCGCCACCATGGAACAAACCGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGCGAAAGATTCACCGTATTTGGCGCCACCATGGAACAAACCGCCAGAAAGGTGAAATCATG 2520
QY 2521 CCGAAACATCCGCGAGATGTCGCTTTCTGGTATGCGCGTGGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAAACATCCGCGAGATGTCGCTTTCTGGTATGCGCGTGGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGGTCTGACAGCTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTCTGACAGCTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACACAAACAAATTAACAATAACAACCTCGGGATCGAGGAGAGATTTCAAAATTC 2700
Db 2641 AACACAAACAAATTAACAATAACAACCTCGGGATCGAGGAGAGATTTCAAAATTC 2700
QY 2701 GGATCTCTTCTCTGTGGCCAGGCGCCCTCGAGCCCGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTGGCCAGGCGCCCTCGAGCCCGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGAATGTGGTAACTCTCTCTCAGAGCTCTCACTGCTGCGCCACCGAGCTACCCAC 2820
Db 2761 CCGAATGTGGTAACTCTCTCTCAGAGCTCTCACTGCTGCGCCACCGAGCTACCCAC 2820
QY 2821 ACSSGTGAACAAACCGTATTAATGCCAGAGTGGGCAAAATCTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACSSGTGAACAAACCGTATTAATGCCAGAGTGGGCAAAATCTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGCTGCGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCAGAAATGTGGC 2940
Db 2881 CTGCTGCGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCAGAAATGTGGC 2940
QY 2941 AAGTCTTTCTCGTGTGCAAAATCTGCTCGGCAACAAAGTACTCACACCGGGAGAAG 3000
Db 2941 AAGTCTTTCTCGTGTGCAAAATCTGCTCGGCAACAAAGTACTCACACCGGGAGAAG 3000
QY 3001 CCTATGTTGTCGGAATGTTGTAAGTCTTTCAGCGCGAGCGATACCTGTTGCGCCAC 3060
Db 3001 CCTATGTTGTCGGAATGTTGTAAGTCTTTCAGCGCGAGCGATACCTGTTGCGCCAC 3060
QY 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAATGCCAGAGTGGCGAAATCTTTTAGT 3120
Db 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAATGCCAGAGTGGCGAAATCTTTTAGT 3120
QY 3121 CAGCGCGCCACCTGGCCAGCCATCAACGCACTCATCTATGCGGAGAGCCATCAAAATGT 3180
Db 3121 CAGCGCGCCACCTGGCCAGCCATCAACGCACTCATCTATGCGGAGAGCCATCAAAATGT 3180
QY 3181 CCAGAAATGTGGCAAGTCTTCTCTCGGTGTGCAAAATCTGTCGCGCACCAACGCTACTC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTCTCTCGGTGTGCAAAATCTGTCGCGCACCAACGCTACTC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCCAGGCGGCGCAGTACCCGTCAGACGTTCCGGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCCAGGCGGCGCAGTACCCGTCAGACGTTCCGGACTACGCT 3300
```

RESULT 6

AAH25681

ID AAH25681 standard; DNA; 8101 BP.

XX

AAH25681;

XX

DT 05-SEP-2001 (first entry)

XX

DE Nucleotide sequence of human Cdc25C phosphatase DNA.

XX Fusion protein; maltose binding protein; MBP; Cdc25C phosphatase;
KW Cdc25B1; Cdc25B2; Cdc25B3; cyclin-dependent kinase; CDK; cell division;
KW ss.
XX
XX Homo sapiens.
XX
XX WO200144467-A2.
XX
XX 21-JUN-2001.
XX
XX 13-DEC-2000; 2000WO-FR003496.
XX
XX 14-DEC-1999; 99FR-00015722.
XX
XX 30-MAY-2000; 2000FR-00006883.
XX
XX 21-SEP-2000; 2000FR-00012008.
XX
XX (SCRS) SCRAS SOC CONSEILS RECH & APPL SCI.
XX
XX Goubin-Gramatica F, Ducommun B, Prevost G;
XX
XX WPI; 2001-398152/42.
XX
XX New fusion protein of maltose-binding protein and Cdc phosphatase, useful
XX for identifying phosphatase modulators for regulating the cell cycle.
XX
XX Claim 7; Page 9-13; 56pp; French.
XX
XX The present sequence encodes human Cdc25C phosphatase, and was expressed
XX in Escherichia coli. It is used to make fusion proteins with Escherichia
XX coli maltose binding protein (MBP). The specification describes fusion
XX proteins of MBP and one of Cdc 25B1, 25B2, 25B3 and 25C. Cdc phosphatases
XX are involved in activation of cyclin-dependent kinases (CDK) that control
XX cell division. The fusion proteins are used to identify modulators of the
XX specified human Cdc phosphatases, potentially useful for regulation of
XX cell division. They are also for studying physiological or
XX physiopathological activities of Cdc phosphatase
XX
XX Sequence 8101 BP; 2127 A; 1987 C; 2121 G; 1866 T; 0 U; 0 Other;
XX
XX Query Match 82.0%; Score 2705.8; DB 4; Length 8101;
XX Best Local Similarity 99.9%; Pred No. 0;
XX Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX 1 CGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGGCATGATAGCGCCGGAAGAGA 60
XX
XX 1 CGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGGCATGATAGCGCCGGAAGAGA 60
XX
XX 61 GTCATTTCAGGTGTGTAATGTGAACCCAGTAACTGATGATGCGGATGCGG 120
XX
XX 61 GTCATTTCAGGTGTGTAATGTGAACCCAGTAACTGATGATGCGGATGCGG 120
XX
XX 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACCCAGCCAGCCAGTTTCGCGAATA 180
XX
XX 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACCCAGCCAGCCAGTTTCGCGAATA 180
XX
XX 181 CGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTTACATCCCAACCCGCTGGCAC 240
XX
XX 181 CGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTTACATCCCAACCCGCTGGCAC 240
XX
XX 241 AACAACTGCGGGGAAAAAGTGTGCTGATTTGGCGTGGCCACCTCCAGTCTGGCCCTGC 300
XX
XX 241 AACAACTGCGGGGAAAAAGTGTGCTGATTTGGCGTGGCCACCTCCAGTCTGGCCCTGC 300
XX
XX 301 ACGCCCGTTCGCAATTTGTCGGCGGATTAATCTTCGCGCGCATCAACTGGGTGCCAGCG 360
XX
XX 301 ACGCCCGTTCGCAATTTGTCGGCGGATTAATCTTCGCGCGCATCAACTGGGTGCCAGCG 360
XX
XX 361 TGGTGGTGTGCAATGTAGAAACGAAGCGCGTGCAGAGCCCTGTAAAGCGGCGTGCACAATC 420
XX
XX 361 TGGTGGTGTGCAATGTAGAAACGAAGCGCGTGCAGAGCCCTGTAAAGCGGCGTGCACAATC 420
XX
XX 421 TTCTCGCGAAACGCGTCAGTGGGCTGATCATTTAACTATCCGCTGGATGACCAAGATGCCA 480

Db
421 TTCTCGCGAAACGCGTCAGTGGGCTGATCATTTAACTATCCGCTGGATGACCAAGATGCCA 480
Qy
481 TTGCTGTGGAAGCTGCCTGCATTAATGTTCCGGCGTATTTCTTCTGATGCTCTGACCCAGA 540
Db
481 TTGCTGTGGAAGCTGCCTGCATTAATGTTCCGGCGTATTTCTTCTGATGCTCTGACCCAGA 540
Qy
541 CACCCATCAACAGATATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Db
541 CACCCATCAACAGATATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Qy
601 TGGTCGCAATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCGG 660
Db
601 TGGTCGCAATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCGG 660
Qy
661 CGCGTCTGCGTCTGCGTGGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db
661 CGCGTCTGCGTCTGCGTGGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy
721 CGGAAACGGAAAGGCGACTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db
721 CGGAAACGGAAAGGCGACTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Qy
781 ATGAGGCGCATGTTCCCACTGGGATGCTGTTGCCAACGATCAGATGGCGCTGGCGGCA 840
Db
781 ATGAGGCGCATGTTCCCACTGGGATGCTGTTGCCAACGATCAGATGGCGCTGGCGGCA 840
Qy
841 TGCGGCGCATTAACCGAGTCCGGCTGCGGCTGGTGGCGGATATCTCGGTAGTGGGATAGC 900
Db
841 TGCGGCGCATTAACCGAGTCCGGCTGCGGCTGGTGGCGGATATCTCGGTAGTGGGATAGC 900
Qy
901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAACCATCAAAACAGATTTTC 960
Db
901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAACCATCAAAACAGATTTTC 960
Qy
961 GCCTGCTGGGCGAAACCGAGCTGGAGCCGCTGCTGCAACTCTCTCAGGCCAGGCGGTGA 1020
Db
961 GCCTGCTGGGCGAAACCGAGCTGGAGCCGCTGCTGCAACTCTCTCAGGCCAGGCGGTGA 1020
Qy
1021 AGGGCAATCAGCTGTTGCGCGTCTCACTGTTGAAAGAAACCAACCCCTGGCGGCCAATA 1080
Db
1021 AGGGCAATCAGCTGTTGCGCGTCTCACTGTTGAAAGAAACCAACCCCTGGCGGCCAATA 1080
Qy
1081 CGCAAAACCGCTCTCCCGCGCTGGCGGATTAATGAGTGTAGTCTACTCTCTAGTCTCTAG 1140
Db
1081 CGCAAAACCGCTCTCCCGCGCTGGCGGATTAATGAGTGTAGTCTACTCTCTAGTCTCTAG 1140
Qy
1141 CCGACTGGAAGCGGCGAGTGAGGCGAAACGCAATTAATGAGTGTAGTCTACTCTCTAGT 1200
Db
1141 CCGACTGGAAGCGGCGAGTGAGGCGAAACGCAATTAATGAGTGTAGTCTACTCTCTAGT 1200
Qy
1201 GCACAAATCTCATGTTTTCAGACGCTTATCATCGACTGCGAGGTGACCAATGCTCTGGCG 1260
Db
1201 GCACAAATCTCATGTTTTCAGACGCTTATCATCGACTGCGAGGTGACCAATGCTCTGGCG 1260
Qy
1261 TCAGCAGCCCATCGGAAGCTGTGGTATGCTGTGAGGTGCGTAAATCACTGATTAATTCG 1320
Db
1261 TCAGCAGCCCATCGGAAGCTGTGGTATGCTGTGAGGTGCGTAAATCACTGATTAATTCG 1320
Qy
1321 TGTGCTCAAGCGGCACTCCCGTTCTGGAATGTTGTTTTTGGCGCGCATCATTAACGGTT 1380
Db
1321 TGTGCTCAAGCGGCACTCCCGTTCTGGAATGTTGTTTTTGGCGCGCATCATTAACGGTT 1380
Qy
1381 CTGGCAAAATTTCTGAAATGAGCTGTTGCAAAATTAATCATCGGCTCGTATTAATGTTGTA 1440
Db
1381 CTGGCAAAATTTCTGAAATGAGCTGTTGCAAAATTAATCATCGGCTCGTATTAATGTTGTA 1440
Qy
1441 ATTGTGAGCGGATAACCAATTTTTCACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCAGA 1500
Db
1441 ATTGTGAGCGGATAACCAATTTTTCACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCAGA 1500
Qy
1501 GCATTTCACAAAGGACCATAGATTATGAAAACTGAAGAGGTAAATCTGGTAAATCTGG 1560
Db
1501 GCATTTCACAAAGGACCATAGATTATGAAAACTGAAGAGGTAAATCTGGTAAATCTGG 1560

Query Match 82.0%; Score 2705.8; DB 6; Length 8101;									
Best Local Similarity 99.9%; Pred. No. 0;									
Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;									
Qy	1	CGACACCATCAATGTGTGCAAAACCTTTTCGGCGGTATGGCATGATAGCGCCCGGAAGAGA	60						
Db	1	CGACACCATCGAATGGTGCAAACCTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA	60						
Qy	61	GTCAATTCAAGGTGTGAATGTGAACCAAGTAACTTATACGATGTCCAGAGTATGCCG	120						
Db	61	GTCAATTCAAGGTGTGAATGTGAACCAAGTAACTTATACGATGTCCAGAGTATGCCG	120						
Qy	121	GTGTCTCTATCAGACCGTTTCCCGCGTGTGAACCGCGCAGCCACGCTTTCGCGAATA	180						
Db	121	GTGTCTCTATCAGACCGTTTCCCGCGTGTGAACCGCGCAGCCACGCTTTCGCGAATA	180						
Qy	181	CGCGGAAAGTGAAGCGCGATGGCGAGCTGAATTAACATCCCAACCGCGTGGCAC	240						
Db	181	CGCGGAAAGTGAAGCGCGATGGCGAGCTGAATTAACATCCCAACCGCGTGGCAC	240						
Qy	241	AACAACTGGCGGCAACAGTCGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300						
Db	241	AACAACTGGCGGCAACAGTCGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300						
Qy	301	AGCGCGCTCGCAAAATTTGTTCGGCGCATTAATCTCTGGCGCGATCAA	360						
Db	301	AGCGCGCTCGCAAAATTTGTTCGGCGCATTAATCTCTGGCGCGATCAA	360						
Qy	361	TGGTGGTGTGATGTGAGAACGAGCGGTGAGAGCGCTGTAAAGCGCGTGCACAATC	420						
Db	361	TGGTGGTGTGATGTGAGAACGAGCGGTGAGAGCGCTGTAAAGCGCGTGCACAATC	420						
Qy	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACCAAGTGC	480						
Db	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACCAAGTGC	480						
Qy	481	TTGCTGTGGAAGCTGCCTGCACTAATGTTCGGCGTATTTCTTGATGTCTGACACGA	540						
Db	481	TTGCTGTGGAAGCTGCCTGCACTAATGTTCGGCGTATTTCTTGATGTCTCTGACACGA	540						
Qy	541	CACCCATCAACAGTATTTCTCCATGAAGACGGTACGGACTGGCGGTGGAGCATC	600						
Db	541	CACCCATCAACAGTATTTCTCCATGAAGACGGTACGGACTGGCGGTGGAGCATC	600						
Qy	601	TGGTGCATTTGGTCAACAGCAAACTCGCTGTAGCGGGCCCAATTAAGTCTGTCTCGG	660						
Db	601	TGGTGCATTTGGTCAACAGCAAACTCGCTGTAGCGGGCCCAATTAAGTCTGTCTCGG	660						
Qy	661	CGCGTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
Db	661	CGCGTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
Qy	721	CGGAACGGGAAGCGACTGGAGTGCATGCTCGGTTTCAACAAACCATGCAATGCTGA	780						
Db	721	CGGAACGGGAAGCGCGACTGGAGTGCATGCTCGGTTTCAACAAACCATGCAATGCTGA	780						
Qy	781	ATGAGGGCATCTTCCCATCGCATGCTGTTTGCACCAACGATCAGATGGCGCTGGCGCAA	840						
Db	781	ATGAGGGCATCTTCCCATCGCATGCTGTTTGCACCAACGATCAGATGGCGCTGGCGCAA	840						
Qy	841	TGCGGCGCAATTACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG	900						
Db	841	TGCGGCGCAATTACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG	900						
Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCATCAAAACAGGATTTTC	960						
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCATCAAAACAGGATTTTC	960						
Qy	961	GCCTGCTGGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGGGTGA	1020						
Db	961	GCCTGCTGGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGGGTGA	1020						
Qy	1021	AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAGAAACCAACCTGGGCGCCAAATA	1080						

Db	1021	AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAGAAACCAACCTGGGCGCCAAATA	1080						
Qy	1081	CGAAAAACCGCTCTCCCGCGCGTTGGCCGATTCATTAAATGACGTGGCACGACAGGTTT	1140						
Db	1081	CGAAAAACCGCTCTCCCGCGCGTTGGCCGATTCATTAAATGACGTGGCACGACAGGTTT	1140						
Qy	1141	CCGACCTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200						
Db	1141	CCGACCTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200						
Qy	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCAACCAATGCTTCTGCGG	1260						
Db	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCAACCAATGCTTCTGCGG	1260						
Qy	1261	TCAGGACGCCATCGGAAGCTGTGTATGGTGTGACGGTCTGTAATCATCTGCATAAATTCG	1320						
Db	1261	TCAGGACGCCATCGGAAGCTGTGTATGGTGTGACGGTCTGTAATCATCTGCATAAATTCG	1320						
Qy	1321	TGTGCTCAAGCGCGACTCCCGTTCTGGATTAATGTTTTTGGCGCGACATCAATACGGTT	1380						
Db	1321	TGTGCTCAAGCGCGACTCCCGTTCTGGATTAATGTTTTTGGCGCGACATCAATACGGTT	1380						
Qy	1381	CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTTGGA	1440						
Db	1381	CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTTGGA	1440						
Qy	1441	ATTGTGAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTGGTGTTCACGA	1500						
Db	1441	ATTGTGAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTGGTGTTCACGA	1500						
Qy	1501	GCATTTCAACCAAGGACCATAGATTATGAAAACTGAAGAAAGGTAAACTGGTAACTCGG	1560						
Db	1501	GCATTTCAACCAAGGACCATAGATTATGAAAACTGAAGAAAGGTAAACTGGTAACTCGG	1560						
Qy	1561	ATTAAACGGCGATAAGGCTATAACGGTCTCCGCTGAGTCCGTTAGAAATTCGAGAAAGAT	1620						
Db	1561	ATTAAACGGCGATAAGGCTATAACGGTCTCCGCTGAGTCCGTTAGAAATTCGAGAAAGAT	1620						
Qy	1621	ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAACTGGAAAGAGAAATTCGCCACAGTT	1680						
Db	1621	ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAACTGGAAAGAGAAATTCGCCACAGTT	1680						
Qy	1681	CGGCAACTGGCGATGGCCCTGACATATCTTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740						
Db	1681	CGGCAACTGGCGATGGCCCTGACATATCTTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740						
Qy	1741	GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGAT	1800						
Db	1741	GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGAT	1800						
Qy	1801	CCGTTTACCTGGGATGCGGTACGTTTCAACCGCAAGCTGATTGCTTACCCGATCCGCTGTT	1860						
Db	1801	CCGTTTACCTGGGATGCGGTACGTTTCAACCGCAAGCTGATTGCTTACCCGATCCGCTGTT	1860						
Qy	1861	GAAGCGTTATCGCTGATTATTAACAAAGATCTGTGCGGAAACCCGCCAAAAACCTGGGAA	1920						
Db	1861	GAAGCGTTATCGCTGATTATTAACAAAGATCTGTGCGGAAACCCGCCAAAAACCTGGGAA	1920						
Qy	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGGAAAGGTAAAGCGCGCTGATGTTCAAC	1980						
Db	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGGAAAGGTAAAGCGCGCTGATGTTCAAC	1980						
Qy	1981	CTGCAAGAACCGTACTCTCACCTGGCGCTGATTGCTGTGACGGGGTTATCGCTTCAAG	2040						
Db	1981	CTGCAAGAACCGTACTCTCACCTGGCGCTGATTGCTGTGACGGGGTTATCGCTTCAAG	2040						
Qy	2041	TATGAAAAACGCAAGTACGACATTAAGACCGTGGGCGTGGATTAACGCTGGCGCGAAAGCG	2100						
Db	2041	TATGAAAAACGCAAGTACGACATTAAGACCGTGGGCGTGGATTAACGCTGGCGCGAAAGCG	2100						
Qy	2101	GGTCTGACCTTCTCGGTTGACCTGATTTAAAAACAAACATGAATTCGAGACACCGATTTAC	2160						

```
Db 2101 GGTCTGACCTTCTCGTGGTGAACCTGATTAATAAAACAAACATGAATGACAGACACCGGATTAAC 2160
QY 2161 TCCATCGCAGAGCTGCTCTTTAATAAGCGCAAAACAGCGATGACCATCAACGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTCTTTAATAAGCGCAAAACAGCGATGACCATCAACGCGCCGCTGG 2220
QY 2221 GCATGGTCAACATCGACACACGCAAGAGTGAATTAATGTTAAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACACGCAAGAGTGAATTAATGTTAAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGTCACCATCCAAACCGTTCTGTCGCTGCTGAGCGAGGTATTAAGCGCCCGAGT 2340
Db 2281 AAGGTCACCATCCAAACCGTTCTGTCGCTGCTGAGCGAGGTATTAAGCGCCCGAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGGTCGT 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGGTCGT 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGATTTGCGCCACCATGGAAACGCCGAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTTGCGCCACCATGGAAACGCCGAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCGCTTTCTGCTGATGCGCTGCTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCGCTTTCTGCTGATGCGCTGCTGCTGCTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCCGACGACTAAATCGAGTCTG 2640
Db 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCCGACGACTAAATCGAGTCTG 2640
QY 2641 AACAAACAAACAAATAACAAATAACAAACAAACCTCGGGATCGAGGGAAGGATTTTCAAGATTC 2700
Db 2641 AACAAACAAACAAATAACAAATAACAAACAAACCTCGGGATCGAGGGAAGGATTTTCAAGATTC 2700
QY 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709

RESULT 8
ADO23608
ID ADO23608 standard; DNA; 7259 BP.
XX
AC ADO23608;
XX
DT 01-JUL-2004 (first entry)
XX
DE DNA encoding MBP-Toxop30del18 (82-294aa) fusion protein.
XX
KW P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; gene.
XX
OS Toxoplasma gondii.
OS Synthetic.
XX
PN US2004067239-A1.
XX
PD 08-APR-2004.
XX
PF 02-OCT-2002; 2002US-00263153.
XX
PR 02-OCT-2002; 2002US-00263153.
XX
PA (MAIN/) MAINE G T.
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX
```

```
PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX
DR WPI: 2004-304563/28.
DR P-PSDB; ADO23609.
XX
PT Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
PT antibodies to Toxoplasma gondii.
XX
PS Example 2; Fig 22; 114pp; English.
XX
CC The invention relates to a purified P30 antigen (I) chosen from 3 fully
CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
CC fusion proteins. (I) is useful for detecting the presence of IgM
CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC ToxoP30 fusion protein of the invention.
XX
SQ Sequence 7259 BP; 1870 A; 1822 C; 1926 G; 1641 T; 0 U; 0 Other;
```

```
Query Match 81.5%; Score 2690.4; DB 12; Length 7259;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCGACACCATCGATGGTGCAGAACCTTTCCGGTATGGCATGATACCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGATGGTGCAGAACCTTTCCGGTATGGCATGATACCGCCCGGAAGAGA 60
QY 61 GTCAAATTCAGGCTGGTGAATGTGAACACAGTATACAGTTATACGATGTGCGAGAGTATCCG 120
Db 61 GTCAAATTCAGGCTGGTGAATGTGAACACAGTATACGATGTGCGAGAGTATCCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCCGGTGTGAACCGAGCCAGCCAGTTCTTCGCGAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGGTGTGAACCGAGCCAGCCAGTTCTTCGCGAAA 180
QY 181 CGCGGAAAAAGTGGAGCGCGATGCGGAGCTGAATTACATTCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGGAGCGCGATGCGGAGCTGAATTACATTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAAAACAGTCTGCTGATTTGGCGTTGCGACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGCAAAACAGTCTGCTGATTTGGCGTTGCGACCTCCAGTCTGGCCCTGC 300
QY 301 ACGGCGCGTGCAGAAATGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCGACGC 360
Db 301 ACGGCGCGTGCAGAAATGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCGACGC 360
QY 361 TGGTGGTGTGATGAGAACGAGCGGCTCGAAGCGCTGAAGCGCGGTGCACAATC 420
Db 361 TGGTGGTGTGATGAGAACGAGCGGCTCGAAGCGCTGAAGCGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACCGTCAAGTGGGCTGATCAATTAATCTCGCTGGATGACGAGATGCCA 480
Db 421 TTCTCGCGCAACCGTCAAGTGGGCTGATCAATTAATCTCGCTGGATGACGAGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAATTTCCGGCGTTATTTCTTGATGTCTTGACCA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAATTTCCGGCGTTATTTCTTGATGTCTTGACCA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGCTAGCGACTGGGGGTGGAGATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGCTAGCGACTGGGGGTGGAGATC 600
```

QY 601 TGGTCGCATTGGGTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
DB 601 TGGTCGCATTGGGTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGCTGCGTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTCGCTGCGTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
DB 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTGGATGCTGGTTGCCAACGATCAGATGCGCTGGCGCGAA 840
DB 781 ATGAGGGCATGTTCCCACTGGATGCTGGTTGCCAACGATCAGATGCGCTGGCGCGAA 840
QY 841 TCGCGGCCATTACCGAGTCCGGGCTCGCGTGGTGGCGGATATCTCGGTAGTGGGATACG 900
DB 841 TCGCGGCCATTACCGAGTCCGGGCTCGCGTGGTGGCGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCCACATCAAAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCCACATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGGCAACACGCGTGGACCGCTTGCCTGCAACTCTCTCAGGSCCAGGCGGTGA 1020
DB 961 GCCTGCTGGGGCAACACGCGTGGACCGCTTGCCTGCAACTCTCTCAGGSCCAGGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTTGCCCGCTCACTGGTGAAGAAACACCCCTGGCGGCCAATA 1080
DB 1021 AGGCAATCAGCTGTTGCCCGCTCACTGGTGAAGAAACACCCCTGGCGGCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCACTGGCGACAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCACTGGCGACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTCATTAG 1200
DB 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTCATTAG 1200
QY 1201 GCACAATCTCATGTTTGACAGCTTATCATGACTGCACGGTGCACCAATGTTCTGGCG 1260
DB 1201 GCACAATCTCATGTTTGACAGCTTATCATGACTGCACGGTGCACCAATGTTCTGGCG 1260
QY 1261 TCAGCAGCCATCGGAAGCTGGTATGCGTGTGAGGTCGTAATCACTGCATAAATTCG 1320
DB 1261 TCAGCAGCCATCGGAAGCTGGTATGCGTGTGAGGTCGTAATCACTGCATAAATTCG 1320
QY 1321 TGTGCTCAGGCGCACTCCCGTTCTGGATATGTTTTTGGCGCGACATCAATACGGTT 1380
DB 1321 TGTGCTCAGGCGCACTCCCGTTCTGGATATGTTTTTGGCGCGACATCAATACGGTT 1380
QY 1381 CTGGCAATATTCTGAAATGAGCTGTTGACAATTAATCATCGGCTCGTATATGTGGA 1440
DB 1381 CTGGCAATATTCTGAAATGAGCTGTTGACAATTAATCATCGGCTCGTATATGTGGA 1440
QY 1441 ATTGTAGCGGATAACAATTTACACAGGAACAGCCAGTCCGTTTAGTGTTTTTCAGCA 1500
DB 1441 ATTGTAGCGGATAACAATTTACACAGGAACAGCCAGTCCGTTTAGTGTTTTTCAGCA 1500
QY 1501 GCATCTTACCACACAGGACCATAGATTATGAAAACTGAAGAAGTTAACTGTAATCTGG 1560
DB 1501 GCATCTTACCACACAGGACCATAGATTATGAAAACTGAAGAAGTTAACTGTAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGCTCTCGCTGAGTCGGTGAAGAAATTCAGAAAGAT 1620
DB 1561 ATTAACGGCGATAAAGGCTATAACGCTCTCGCTGAGTCGGTGAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCAATCGGCGATGGGCCCTGACATTATCTCTGCGGCAACAGCGCTTTGGTGGCTAC 1740

RESULT 9
ADO23613
ID ADO23613 standard; DNA; 7322 BP.
XX
AC ADO23613;

DB 1681 GCGGCAATCGGCGATGGCCCTGACATTATCTTCTGGGCAACAGCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCAACCCCGSACAAGCGTTCCAGGACAAGCTGAT 1800
DB 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCAACCCCGSACAAGCGTTCCAGGACAAGCTGAT 1800
QY 1801 CCGTTTACCTCGGATGCGTACGTTACAACGGCAAGCTGATTGTTACCCGATCGCTGTT 1860
DB 1801 CCGTTTACCTCGGATGCGTACGTTACAACGGCAAGCTGATTGTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
DB 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAACGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
DB 1921 GAGATCCCGCGCTGGATTAAGAACTGAAACGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
DB 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
DB 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGTTGACCTGATTAAAAACAACATGAAATGCGAGACACCGATTAC 2160
DB 2101 GGTCTGACCTTCTGTTGACCTGATTAAAAACAACATGAAATGCGAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGTGCCTTTAATAAGGCGAACAAGCGATGACCATCAAGCGGCCGTGG 2220
DB 2161 TCCATCGCAGAGTGCCTTTAATAAGGCGAACAAGCGATGACCATCAAGCGGCCGTGG 2220
QY 2221 GCATGTCACACATCGACACACAGCAAAAGTGAATTAATGTTGTAACCGTACTGCCGACCTTC 2280
DB 2221 GCATGTCACACATCGACACACAGCAAAAGTGAATTAATGTTGTAACCGTACTGCCGACCTTC 2280
QY 2281 AAGGTCACACATCCAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAAACGCCCCAGT 2340
DB 2281 AAGGTCACACATCCAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAAACGCCCCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
DB 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGTTAATAAGACAAACCGCTGGTGCCTGAGCGCTGAAGTCTTACGAGGAAGAG 2460
DB 2401 GAAGCGTTAATAAGACAAACCGCTGGTGCCTGAGCGCTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGCAAGAGATCCAAGTATTGCGGCCACCATGGAACAAACGCCAGAAAGGTGAATCATG 2520
DB 2461 TTGGCAAGAGATCCAAGTATTGCGGCCACCATGGAACAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACAATCCCGCAGATGTCGCTTTCTGGTATGTCGCTGCGTATCTGCGGTGATCAACGCC 2580
DB 2521 CCGAACAATCCCGCAGATGTCGCTTTCTGGTATGTCGCTGCGTATCTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACGCTGTCGAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
DB 2581 GCCAGCGTCTGACGCTGTCGAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACACAAACAATAAACAACAAACCTCGGGATCGAGGGAAGGATTTCAGAAATTC 2700
DB 2641 AACACAAACAATAAACAACAAACCTCGGGATCGAGGGAAGGATTTCAGAAATTC 2700

XX	01-JUL-2004	(first entry)
DT		
XX		
DE	DNA encoding MBP-ToxoP30del10(52-284aa) fusion protein.	
XX		
KW	P30 antigen; Toxo30del13C; Toxo30del2C; ToxoP30 MIX1;	
KW	MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;	
KW	Toxoplasma gondii; db; gene.	
XX		
OS	Toxoplasma gondii.	
OS	Synthetic.	
XX		
PN	US2004067239-A1.	
XX		
PD	08-APR-2004.	
XX		
XX	02-OCT-2002; 2002US-00263153.	
PF		
XX		
PR	02-OCT-2002; 2002US-00263153.	
XX		
PA	(MAIN/) MAINE G T.	
PA	(PATE/) PATEL C B.	
PA	(GINS/) GINSBURG S R.	
PA	(BLIE/) BLIESE T R.	
XX		
PI	Maine GT, Patel CB, Ginsburg SR, Bliese TR;	
XX		
DR	WPI; 2004-304563/28.	
DR	P-PSDB; ADO23614.	
XX		
PT	Novel purified polypeptide having sequence identity to amino acid	
PT	sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,	
PT	Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG	
PT	antibodies to Toxoplasma gondii.	
XX		
FS	Example 2; Fig 25; 114pp; English.	
XX		
CC	The invention relates to a purified P30 antigen (I) chosen from 3 fully	
CC	defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6	
CC	amino acids added to the C-terminus of the amino acid sequence of	
CC	Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in	
CC	which at least one of the five C-terminal cysteine amino acids of the	
CC	amino acid sequence of Toxo30del13C P30 antigen sequence is substituted	
CC	with alanine, or comprising the amino acid sequence chosen from MBP-	
CC	Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1	
CC	fusion proteins. (I) is useful for detecting the presence of IgM	
CC	antibodies to Toxoplasma gondii in a test sample, which involves	
CC	contacting the test sample suspected of containing the IgM antibodies	
CC	with a composition comprising (I) and detecting the presence of (I)/IgM	
CC	antibody complexes. The present sequence represents DNA encoding a MBP-	
CC	ToxoP30 fusion protein of the invention.	
XX		
SQ	Sequence 7322 BP; 1887 A; 1844 C; 1934 G; 1657 T; 0 U; 0 Other;	
Query Match	81.5%; Score 2590.4; DB 12; Length 7322;	
Best Local Similarity	99.8%; Pred. No. 0;	
Matches 2694; Conservative	0; Mismatches 6; Indels 0; Gaps 0	
Qy	1 CCGACACCATCGAATGGTGC AAAACCTTTTCGGGTATGGCATGATAGCGCCGGGAAGA 60	
Db	1 CCGACACCATCGAATGGTGC AAAACCTTTTCGGGTATGGCATGATAGCGCCGGNAGAGA 60	
Qy	61 GTCAATTCAGGGTGGTGAATGTGAAACCAAGTAAAGTTATACGATGTCGAGAGTATGCCG 120	
Db	61 GTCAATTCAGGGTGGTGAATGTGAAACCAAGTAAAGTTATACGATGTCGAGAGTATGCCG 120	
Qy	121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGCGCCAGCCAGCTTTCTCGGAAA 180	
Db	121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGCGCCAGCCAGCTTTCTCGGAAA 180	
Qy	181 CGCGGGAAAAGTGGAGCGCGCATGGCGGAGCTGAATTATCATTTCCCAACCGCTGGCAC 240	
Db	181 CGCGGGAAAAGTGGAGCGCGCATGGCGGAGCTGAATTATCATTTCCCAACCGCTGGCAC 240	

QY 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTTGGCCGCAATCATATAACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTTGGCCGCAATCATATAACGGTT 1380
QY 1381 CTGGCAAAATATTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGTGA 1440
Db 1381 CTGGCAAAATATTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGTGA 1440
QY 1441 ATTGTGAGCGGATAACAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCAACA 1500
Db 1441 ATTGTGAGCGGATAACAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCAACA 1500
QY 1501 GCACCTTACCAACAGGACCATAGATTATGAAAACCTGAAGAAGTAAACTGTTATCTGG 1560
Db 1501 GCACCTTACCAACAGGACCATAGATTATGAAAACCTGAAGAAGTAAACTGTTATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCCGTTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCCGTTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAAGTGAAGAGAAAATTTCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAAGTGAAGAGAAAATTTCCACAGGTT 1680
QY 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGGCTGTGGCTGAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGGCTGTGGCTGAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
QY 1801 CGGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATGCTTTACCCGATCGCTGT 1860
Db 1801 CGGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATGCTTTACCCGATCGCTGT 1860
QY 1861 GAAGGCTTATCGCTGATTTAACAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGGCTTATCGCTGATTTAACAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTCGCGCGCTGATGCTGACGGGGGTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTCGCGCGCTGATGCTGACGGGGGTTATCGGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGTTGGGGGTGGATAACGCTGGCGGCAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGTTGGGGGTGGATAACGCTGGCGGCAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAACCAACACATCAATGACGACACCGGATAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAACCAACACATCAATGACGACACCGGATAC 2160
QY 2161 TCCATCGCAGAGCTGCGCTTTAATAAGCGGAAACAGCGATCACCATCAACCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCGCTTTAATAAGCGGAAACAGCGATCACCATCAACCGCCGCTGG 2220
QY 2221 GCATGGTCCAACTCGACACAGCAAGAGTGAATTAATGTTGTAACGTAACGCGACCTTC 2280
Db 2221 GCATGGTCCAACTCGACACAGCAAGAGTGAATTAATGTTGTAACGTAACGCGACCTTC 2280
QY 2281 AAGGCTCAACCATCAACCGTTCTGTTGGCGTCTGAGGCGAGGTATTAAACCGCCCAAGT 2340
Db 2281 AAGGCTCAACCATCAACCGTTCTGTTGGCGTCTGAGGCGAGGTATTAAACCGCCCAAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACATCTCTGCTGACTGATCAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACATCTCTGCTGACTGATCAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGCTAGCGCTGAAGTCTTTACGAGGAAGAG 2460

Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGCTGAAAGTCTTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCACGTTATTGCGCGCCACCATGGAACACGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATTGCGCGCCACCATGGAACACGCCAGAAAGGTGAAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCGTGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCGTGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGGTGCTGACACTGTCGATGAAGCCCTGAAAGACGCCGACACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGCTGACACTGTCGATGAAGCCCTGAAAGACGCCGACACTAAATTCGAGCTCG 2640
QY 2641 AACCAACAAACAATAACAATAACAACCAACCAACCTCGGGATCGAGGGAAGGATTTCAAAATTC 2700
Db 2641 AACCAACAAACAATAACAATAACAACCAACCAACCTCGGGATCGAGGGAAGGATTTCAAAATTC 2700

RESULT 10
ADO23603
ID ADO23603 standard; DNA; 7352 BP.
XX ADO23603;
XX 01-JUL-2004 (first entry)
XX DNA encoding MBP-ToxoP30del14C(52-294aa) fusion protein.
XX P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; gene.
XX Toxoplasma gondii.
OS Synthetic.
XX US2004067239-A1.
XX 08-APR-2004.
XX 02-OCT-2002; 2002US-00263153.
XX 02-OCT-2002; 2002US-00263153.
XX (MAIN/) MAINE G T.
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX WPI; 2004-304563/28.
DR P-PSDB; ADO23604.
XX Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
PT antibodies to Toxoplasma gondii.
XX Example 2; Fig 19; 114pp; English.
PS The invention relates to a purified P30 antigen (I) chosen from 3 fully
CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
CC fusion proteins. (I) is useful for detecting the presence of IgM
CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies

CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC Toxob30 fusion protein of the invention.
XX
SQ Sequence 7352 BP; 1896 A; 1852 C; 1943 G; 1661 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7352;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY	1	CGACACCATCGAATGGTGCAGAAACCTTTCGGGGTATGCGCATGATAGCCCGGAGAGA	60
DB	1	CGACACCATCGAATGGTGCAGAAACCTTTCGGGGTATGCGCATGATAGCCCGGAGAGA	60
QY	61	GTCAATTACAGGTTGGTGAATGTGAACACAGTAACAGTTATACCATGTTCGAGAGTATGCG	120
DB	61	GTCAATTACAGGTTGGTGAATGTGAACACAGTAACAGTTATACCATGTTCGAGAGTATGCG	120
QY	121	GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCACGTTTCTGCGAAAA	180
DB	121	GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCACGTTTCTGCGAAAA	180
QY	181	CGCGGAAAACTGGAAGCGGATGCGGAGCTGAAATACATTCGCCAACCGCGTGGCAC	240
DB	181	CGCGGAAAACTGGAAGCGGATGCGGAGCTGAAATACATTCGCCAACCGCGTGGCAC	240
QY	241	AACAACTGGCGGCAAAACAGTCGTTGCTGATGGGTTGCCACCTCCAGTCTGGGCCCTGC	300
DB	241	AACAACTGGCGGCAAAACAGTCGTTGCTGATGGGTTGCCACCTCCAGTCTGGGCCCTGC	300
QY	301	ACGCGCGTGCAGAAATTTGTCGGCGGATTAATCTCGCGCGGATCAAATGGGTGCCAGC	360
DB	301	ACGCGCGTGCAGAAATTTGTCGGCGGATTAATCTCGCGCGGATCAAATGGGTGCCAGC	360
QY	361	TGGTGTGTCGATGTAGAACGAGCGGCTGAAAGCCTGTAAGCGGCTGACCAATC	420
DB	361	TGGTGTGTCGATGTAGAACGAGCGGCTGAAAGCCTGTAAGCGGCTGACCAATC	420
QY	421	TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTCCGTCGATGACCAAGATGCCA	480
DB	421	TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTCCGTCGATGACCAAGATGCCA	480
QY	481	TTGCTGTGGAAGTCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTGACGAGA	540
DB	481	TTGCTGTGGAAGTCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTGACGAGA	540
QY	541	CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC	600
QY	601	TGGTGGCATTTGGTTCACAGCAAAATCGCGTGTAGCGGGCCCATTAAGTTCTGTCGG	660
DB	601	TGGTGGCATTTGGTTCACAGCAAAATCGCGTGTAGCGGGCCCATTAAGTTCTGTCGG	660
QY	661	CGCGTCTCGCTGCGTGGCATTAATCTCACTCGCAATCAAAATTCAGCGGATAG	720
DB	661	CGCGTCTCGCTGCGTGGCATTAATCTCACTCGCAATCAAAATTCAGCGGATAG	720
QY	721	CGGAAACGGGAAGGCGACTGGAGTGCATGTCCGGTTTTTCAACAAACCAATGCAATGCTGA	780
DB	721	CGGAAACGGGAAGGCGACTGGAGTGCATGTCCGGTTTTTCAACAAACCAATGCAATGCTGA	780
QY	781	ATGAGGGCATCTGTTCCCACTCGGATGTCGTTGGTCCAGATCAGATGCGCGTGGCGCAA	840
DB	781	ATGAGGGCATCTGTTCCCACTCGGATGTCGTTGGTCCAGATCAGATGCGCGTGGCGCAA	840
QY	841	TGCGGCCATTACAGGTCCGGGTGCGGTTGTCGGATATCTCGGTAGTGGGATACG	900
DB	841	TGCGGCCATTACAGGTCCGGGTGCGGTTGTCGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAAGACAGTCTATGTTTATATCCCGCGTTTAAACCAATCAACAGGATTTTC	960
DB	901	ACGATACCGAAGACAGTCTATGTTTATATCCCGCGTTTAAACCAATCAACAGGATTTTC	960

QY	961	GCCTGCTGGGCAAAACAGCGTGCACCGCTTGTGCAACTCTCTCAGGCCAGCGGTGA	1020
DB	961	GCCTGCTGGGCAAAACAGCGTGCACCGCTTGTGCAACTCTCTCAGGCCAGCGGTGA	1020
QY	1021	AGGCAATCAGCTGTGCTCCGCTCTCAGTGTGAAAGAAAACACCCCTGGCGCCCAATA	1080
DB	1021	AGGCAATCAGCTGTGCTCCGCTCTCAGTGTGAAAGAAAACACCCCTGGCGCCCAATA	1080
QY	1081	CGCAAAACCGCTCTCCCGCGCGTGGCCGATTCAATTAATGCAGCTGGCAGCAGAGTTT	1140
DB	1081	CGCAAAACCGCTCTCCCGCGCGTGGCCGATTCAATTAATGCAGCTGGCAGCAGAGTTT	1140
QY	1141	CCGCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCATCTATTAG	1200
DB	1141	CCGCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCATCTATTAG	1200
QY	1201	GCACAAATCTCATGTTTGCAGCTTATCATCGACTGCAGCTGCACCAATGCTTCTGGCG	1260
DB	1201	GCACAAATCTCATGTTTGCAGCTTATCATCGACTGCAGCTGCACCAATGCTTCTGGCG	1260
QY	1261	TCAGGAGCCATCGAAAGCTGTGATGTGCTGTGAGTGTGCTGTAATCACTGCATAAATTCG	1320
DB	1261	TCAGGAGCCATCGAAAGCTGTGATGTGCTGTGAGTGTGCTGTAATCACTGCATAAATTCG	1320
QY	1321	TGTCGCTCAAGCGGCACTCCCGTCTCGGTAATGTTTTTTCGCGCGAGCATCAAAACGGTT	1380
DB	1321	TGTCGCTCAAGCGGCACTCCCGTCTCGGTAATGTTTTTTCGCGCGAGCATCAAAACGGTT	1380
QY	1381	CTGGCAAAATTTCTGAAATGAGCTGTGTGACAAATTAATCATCGCTGTGTAATGTGTGGA	1440
DB	1381	CTGGCAAAATTTCTGAAATGAGCTGTGTGACAAATTAATCATCGCTGTGTAATGTGTGGA	1440
QY	1441	ATTGTGAGCGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTAGTGTGTTTTCACGA	1500
DB	1441	ATTGTGAGCGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTAGTGTGTTTTCACGA	1500
QY	1501	GCATTCACCAACAAAGGACCATAGATTATGAAACTGAAAGGTAACCTGTTAATCTGG	1560
DB	1501	GCATTCACCAACAAAGGACCATAGATTATGAAACTGAAAGGTAACCTGTTAATCTGG	1560
QY	1561	ATTAACGGCGATAAAGGCTATAACCGTCTCGCTGAACTGAGTAAAGAAATTCGAGAAAGAT	1620
DB	1561	ATTAACGGCGATAAAGGCTATAACCGTCTCGCTGAACTGAGTAAAGAAATTCGAGAAAGAT	1620
QY	1621	ACGGCAATTAAGCTACCGTTGAGCATCCGGATAAAGTAAAGGAGAAATTCACAGGTT	1680
DB	1621	ACGGCAATTAAGCTACCGTTGAGCATCCGGATAAAGTAAAGGAGAAATTCACAGGTT	1680
QY	1681	CGGCAACTGGCGATGGCCCTGACATTATCTCTGGGCAACGACCGCTTGGTGGCTAC	1740
DB	1681	CGGCAACTGGCGATGGCCCTGACATTATCTCTGGGCAACGACCGCTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCCCTGTTGGCTGAAATCAACCGGAAACAAAGCGTTCCAGGCAAGCTGTAT	1800
DB	1741	GCTCAATCTGGCCCTGTTGGCTGAAATCAACCGGAAACAAAGCGTTCCAGGCAAGCTGTAT	1800
QY	1801	CCGTTTACCTGGGATCGCTGACGTTACAGCGCAAGCTGATTTGCTTACCGGATCGCTGTT	1860
DB	1801	CCGTTTACCTGGGATCGCTGACGTTACAGCGCAAGCTGATTTGCTTACCGGATCGCTGTT	1860
QY	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCAAACCCGCAAAACCTCGGAA	1920
DB	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCAAACCCGCAAAACCTCGGAA	1920
QY	1921	GAGATCCCGCGCTGATAAGAACTGAAAGGTAAGGTAAGAGCGCGCTGTATGTTCAAC	1980
DB	1921	GAGATCCCGCGCTGATAAGAACTGAAAGGTAAGGTAAGAGCGCGCTGTATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATTTGCTGTCGACGGGGTTATGCGTTCAAG	2040
DB	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATTTGCTGTCGACGGGGTTATGCGTTCAAG	2040

```
Qy 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATAAACGCTGGCGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATAAACGCTGGCGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGTTGACCTGATTAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAAACAAACACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTTAATTAAGCGGAAACAGCGATGACCATCAACGCGCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAATTAAGCGGAAACAGCGATGACCATCAACGCGCCGTGG 2220
Qy 2221 GCATGGTCAACATCGACACACGAGAAAGTGAATTAATGCTGTAAACGCTACTGCGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACACGAGAAAGTGAATTAATGCTGTAAACGCTACTGCGACCTTC 2280
Qy 2281 AAGGGTCAACCATCAACACCGTTCTGCTGCTGAGCGAGGATTAAACGCGCCCAAGT 2340
Db 2281 AAGGGTCAACCATCAACACCGTTCTGCTGCTGAGCGAGGATTAAACGCGCCCAAGT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGGTAAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTAAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGGCGAAGATCCACGCTATTGCGCGCCACCATGGAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGCTATTGCGCGCCACCATGGAAACGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTGATGCGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTGATGCGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Qy 2581 GCCAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640
Db 2581 GCCAGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640
Qy 2641 AACACACACAAATTAACAAATTAACAAATTAACAAATTAACAAATTAACAAATTAACAAAT 2700
Db 2641 AACACACACAAATTAACAAATTAACAAATTAACAAATTAACAAATTAACAAATTAACAAAT 2700

RESULT 11
ADO23598
ID ADO23598 standard; DNA; 7370 BP.
AC ADO23598;
XX ADO23598;
XX 01-JUL-2004 (first entry)
DT DNA encoding MBP-ToxoP30del13C(52-300aa) fusion protein.
DE P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; db; gene.
XX
OS Toxoplasma gondii.
OS Synthetic.
XX
PN US2004067239-A1.
XX
XX 08-APR-2004.
XX
PF 02-OCT-2002; 2002US-00263153.
PR 02-OCT-2002; 2002US-00263153.
XX (MAIN/) MAINE G T.
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
```

```
PA (BLIE/) BLIESE T R.
XX Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX WPI; 2004-304563/28.
DR P-PSDB; ADO23599.
XX
PT Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
PT antibodies to Toxoplasma gondii.
XX
PS Example 2; Fig 15; 114pp; English.
XX
CC The invention relates to a purified P30 antigen (I) chosen from 3 fully
CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
CC fusion proteins. (I) is useful for detecting the presence of IgM
CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC ToxoP30 fusion protein of the invention.
XX
SQ Sequence 7370 BP; 1901 A; 1855 C; 1949 G; 1665 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGCAGAAACCTTTCGGGTATGGCATGTAGCGCCGGAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAGAAACCTTTCGGGTATGGCATGTAGCGCCGGAGAGA 60
Qy 61 GTCATTCAGGGTGTGAATGTGAACCAAGTATAGCATGTCCGACAGTATGCG 120
Db 61 GTCATTCAGGGTGTGAATGTGAACCAAGTATAGCATGTCCGACAGTATGCG 120
Qy 121 GTGCTCTTATCAGACCGTTTCCCGCGTGTGAACCAAGCGCAGCCACGTTTCTGCGAAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCCGCGTGTGAACCAAGCGCAGCCACGTTTCTGCGAAAA 180
Qy 181 CGCGGAAAAAGTGAACGCGGATGGCGGAGCTGAATTACATTCGCAACCGCTGGCAC 240
Db 181 CGCGGAAAAAGTGAACGCGGATGGCGGAGCTGAATTACATTCGCAACCGCTGGCAC 240
Qy 241 AACAACTGGCGGCAACACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAACACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGCGCGCTCGCAAAATTTCTCGCGCGATTAAATCTCGGCCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGCTCGCAAAATTTCTCGCGCGATTAAATCTCGGCCGATCAACTGGGTGCCAGCG 360
Qy 361 TGGTGGTGTGATGTGAACGAGCGCGCTGCAAGCCTGTAAAGCGCGGTGCAACATC 420
Db 361 TGGTGGTGTGATGTGAACGAGCGCGCTGCAAGCCTGTAAAGCGCGGTGCAACATC 420
Qy 421 TTCTCGCGCAACGCGTCAGTGGCTGATTAATTAATTCGCTGGATGATCACCAGGATGCCA 480
Db 421 TTCTCGCGCAACGCGTCAGTGGCTGATTAATTAATTCGCTGGATGATCACCAGGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCGAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCGAGA 540
Qy 541 CACCACATCAACAGTATTATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCACATCAACAGTATTATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
```

Db 541 CACCCATCAACAGTAATATTTTCTCCATGAAGACGGTACGCGACTGGCGGTGAGCATC 600
QY 601 TGGTCGCATTGGTCAACAGCAAAATCGCGCTGTGTAGCGGCCCAATTAAGTTCTGTCTCGG 660
Db 601 TGGTCGCATTGGTCAACAGCAAAATCGCGCTGTGTAGCGGCCCAATTAAGTTCTGTCTCGG 660
QY 661 CCGGCTCTGGCTCTGGCTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CCGGCTCTGGCTCTGGCTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CCGAAGCGGAAGCGCATGAGTGCATGTCCGGTTTCAACAAACCAATGCAAAATGCTGA 780
Db 721 CCGAAGCGGAAGCGCATGAGTGCATGTCCGGTTTCAACAAACCAATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCACCTGCGATGTGGTTGCCAAGCATCAGATGCGCTGGCGCGCA 840
Db 781 ATGAGGGCATCGTTCACCTGCGATGTGGTTGCCAAGCATCAGATGCGCTGGCGCGCA 840
QY 841 TGGCGGCCAATACCGAGTCCGGCTCGCGTTGGTGGGATATCTCGGTAGTGGGATAG 900
Db 841 TGGCGGCCAATACCGAGTCCGGCTCGCGTTGGTGGGATATCTCGGTAGTGGGATAG 900
QY 901 AGGATACCGAAGCAGCATGTTATATCCCGCGTTAACACCAATCAAAAGGATTTTC 960
Db 901 AGGATACCGAAGCAGCATGTTATATCCCGCGTTAACACCAATCAAAAGGATTTTC 960
QY 961 GCCTCTGGGGCAACACGAGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTCTGGGGCAACACGAGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGGCGGTCTCACTGGTGAAGAAAAACCAACCTGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCGGTCTCACTGGTGAAGAAAAACCAACCTGCGGCCAATA 1080
QY 1081 CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTAATGCAAGTGCAGCAGAGTTT 1140
Db 1081 CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTAATGCAAGTGCAGCAGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAATTCATGTTTGAAGCTTATCATGCACTGCAAGTGCACCAATCTCTGCGC 1260
Db 1201 GCACAATTCATGTTTGAAGCTTATCATGCACTGCAAGTGCACCAATCTCTGCGC 1260
QY 1261 TCAGGACGCAATCGGAAGCTGTGGTATGCGTGTGCGAGTGCATCACTGCAATTTTC 1320
Db 1261 TCAGGACGCAATCGGAAGCTGTGGTATGCGTGTGCGAGTGCATCACTGCAATTTTC 1320
QY 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTTTCGCGGACATCATACCGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTTTCGCGGACATCATACCGTT 1380
QY 1381 CTGGCAATATCTGAATGAGCTGTGCAATTAATCATCGGCTCGTATTAATGTGTGA 1440
Db 1381 CTGGCAATATCTGAATGAGCTGTGCAATTAATCATCGGCTCGTATTAATGTGTGA 1440
QY 1441 ATTGTAGCGGATAACAATTTTCAACAGGAACAGCAGTCCGTTTATAGTGTGTTTCACGA 1500
Db 1441 ATTGTAGCGGATAACAATTTTCAACAGGAACAGCAGTCCGTTTATAGTGTGTTTCACGA 1500
QY 1501 GCATTTCAACCAAGGACATAGCATATGAAATTCGAAGAGGTAAATCTGGTAATCTGG 1560
Db 1501 GCATTTCAACCAAGGACATAGCATATGAAATTCGAAGAGGTAAATCTGGTAATCTGG 1560
QY 1561 ATTAACGCGATAAGGCTATAACGGTCTCGCTGAGTGCATGAGTAAGTAATTCGAGAAAGAT 1620
Db 1561 ATTAACGCGATAAGGCTATAACGGTCTCGCTGAGTGCATGAGTAAGTAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTTCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTTCCACAGGTT 1680

QY 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACAGACCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACAGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCCTTTGGCTGAAATCACCCGCAAAAGGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCCTTTGGCTGAAATCACCCGCAAAAGGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACGCGCAAGCTGATGTTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACGCGCAAGCTGATGTTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAAACGTTACTTCACTGCGCTGATGCTGCTGCAAGGGGTTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAAACGTTACTTCACTGCGCTGATGCTGCTGCAAGGGGTTTATGCGTTCAAG 2040
QY 2041 TATGAAAACGCGAAGTACGACATTAAGACGTTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGCGAAGTACGACATTAAGACGTTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAATAAAACACACATGATGACACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAATAAAACACACATGATGACACACCGATTAC 2160
QY 2161 TCCATCCGAGAAAGTGCCTTTTAAAGGCGAAACAGCGATGACCAATCAACGCGCCGTGG 2220
Db 2161 TCCATCCGAGAAAGTGCCTTTTAAAGGCGAAACAGCGATGACCAATCAACGCGCCGTGG 2220
QY 2221 GCATGTCCAAACATCGACACACGCAAGTGAATTTATGTTGTTAAACGCTGCGGACCTTC 2280
Db 2221 GCATGTCCAAACATCGACACACGCAAGTGAATTTATGTTGTTAAACGCTGCGGACCTTC 2280
QY 2281 AAGGTCACCACTTCAAAACCGTTTCTGTTGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGTCACCACTTCAAAACCGTTTCTGTTGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGATTTCTCGAATACTATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGATTTCTCGAATACTATCTGCTGACTGATGAAGGTCTG 2400
QY 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGGAAAGATCAAGTATTTGCGGCAACATGGAAGAACCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGGAAAGATCAAGTATTTGCGGCAACATGGAAGAACCCAGAAAGGTGAAATCATG 2520
QY 2521 CCGAACAATCCGAGATGTCGCTTTCTGTTATGCGCTGCGTACTGCGGTGATCAAGCC 2580
Db 2521 CCGAACAATCCGAGATGTCGCTTTCTGTTATGCGCTGCGTACTGCGGTGATCAAGCC 2580
QY 2581 GCCAGCGGTCTGTCAGACTGTGATGAAGCCCTGAAAGAGCGCGAGACTTAATTCGAGTCTG 2640
Db 2581 GCCAGCGGTCTGTCAGACTGTGATGAAGCCCTGAAAGAGCGCGAGACTTAATTCGAGTCTG 2640
QY 2641 AACAAACAACAAATTAACAAATTAACAAACCTCGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACAAACAACAAATTAACAAATTAACAAACCTCGGATCGAGGGAAGGATTTTCAGAAATTC 2700

RESULT 12

AD023639

ID AD023639 standard; DNA; 7370 BP.

```
XX AC ADO23639;
XX DT 01-JUL-2004 (first entry)
XX DE DNA encoding MBP-ToxoP30MIX1 fusion protein.
XX KW P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; de; gene.
XX OS Toxoplasma gondii.
XX OS Synthetic.
XX PN US2004067239-A1.
XX PD 08-APR-2004.
XX PF 02-OCT-2002; 2002US-00263153.
XX PR 02-OCT-2002; 2002US-00263153.
XX PA (MAIN/) MAINE G T.
XX PA (PATE/) PATEL C B.
XX PA (GINS/) GINSBURG S R.
XX PA (BLIE/) BLIESE T R.
XX PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX DR WPI; 2004-304563/28.
XX DR P-PSDB; ADO23640.
XX PT Novel purified polypeptide having sequence identity to amino acid
XX PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
XX PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
XX PT antibodies to Toxoplasma gondii.
XX PS Example 5; Fig 31; 114pp; English.
XX CC The invention relates to a purified P30 antigen (I) chosen from 3 fully
XX CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
XX CC amino acids added to the C-terminus of the amino acid sequence of
XX CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
XX CC which at least one of the five C-terminal cysteine amino acids of the
XX CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
XX CC with alanine, or comprising the amino acid sequence chosen from MBP-
XX CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
XX CC fusion proteins. (I) is useful for detecting the presence of IgM
XX CC antibodies to Toxoplasma gondii in a test sample, which involves
XX CC contacting the test sample suspected of containing the IgM antibodies
XX CC with a composition comprising (I) and detecting the presence of (I)/IgM
XX CC antibody complexes. The present sequence represents DNA encoding a MBP-
XX CC ToxoP30 fusion protein of the invention.
XX SQ Sequence 7370 BP; 1900 A; 1856 C; 1950 G; 1664 T; 0 U; 0 Other;
Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCCCGGGAAGAGA 60
DB 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCCCGGGAAGAGA 60
QY 61 GTCATTCCAGGTGGTGAATGTCGAACCACTTATACGTTATACGTTATCGCAGAGTATGCC 120
DB 61 GTCATTCCAGGTGGTGAATGTCGAACCACTTATACGTTATACGTTATCGCAGAGTATGCC 120
QY 121 GTGTCTCTTATCAGACCGGTTTCCCGGTGGTGAACCCAGGCCGACGACGTTTCTGCGAAAA 180
DB 121 GTGTCTCTTATCAGACCGGTTTCCCGGTGGTGAACCCAGGCCGACGACGTTTCTGCGAAAA 180
QY 181 CGCGGGAAAAAGTGGGAAGCGCGCATGGCGGAGCTGAATTCATATCCCAACCCGCTGGCAC 240
```

```
Db 181 CGCGGGAAAAAGTGGGAAGCGCGCATGGCGGAGCTGAATTCATATCCCAACCCGCTGGCAC 240
QY 241 AACAACTGGCGGGCAAAACAGTCGTTGTCTGATTGGCGTTGGCCACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGGCAAAACAGTCGTTGTCTGATTGGCGTTGGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCCGCGTGCAGAAATTTGTCGCGCGGATTAATCTCGCGCGGATCAATCTGGGTGGCAGCG 360
DB 301 ACGCCGCGTGCAGAAATTTGTCGCGCGGATTAATCTCGCGCGGATCAATCTGGGTGGCAGCG 360
QY 361 TGGTGGTGTGATGATGAGAACGAGCGCGCTGGAAGCCTGTAAGCGCGCGTGCACAATC 420
DB 361 TGGTGGTGTGATGATGAGAACGAGCGCGCTGGAAGCCTGTAAGCGCGCGTGCACAATC 420
QY 421 TTCTCGCGCAACGGCTCAGTGGGCTGATCATTAACTATTCCTGATGATGCCAGATGCCA 480
DB 421 TTCTCGCGCAACGGCTCAGTGGGCTGATCATTAACTATTCCTGATGATGCCAGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGGCTTATTTCTTGTGATGCTCTGACCCAGA 540
DB 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGGCTTATTTCTTGTGATGCTCTGACCCAGA 540
QY 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGCTACGCGCTGGCGCTGGAGCATC 600
DB 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGCTACGCGCTGGCGCTGGAGCATC 600
QY 601 TGGTTCGATTGGGTCACCGCAAAATTCGCGCTGTTTAGCGGGGCCAATTAAGTTCTGCTCGG 660
DB 601 TGGTTCGATTGGGTCACCGCAAAATTCGCGCTGTTTAGCGGGGCCAATTAAGTTCTGCTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGGCTGAGCAATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGCGTCTGGCTGGGCTGAGCAATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAAACGGGAAGCGACGTGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
DB 721 CGGAAACGGGAAGCGACGTGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGGATCGTTCCTCCACTCGATGCTGTTGTCGCAACGATCAGATGGCGTGGCGCAAA 840
DB 781 ATGAGGGGATCGTTCCTCCACTCGATGCTGTTGTCGCAACGATCAGATGGCGTGGCGCAAA 840
QY 841 TGGCGCCCATTAACGAGTCCGGGCTGCGGTTGTTGTCGCGATATCTCGGTAGTGGGATAG 900
DB 841 TGGCGCCCATTAACGAGTCCGGGCTGCGGTTGTTGTCGCGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAAGGATTTTC 960
QY 961 GCCTGCTGGGGCAAAACCAACGAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCCAGCGGTGA 1020
DB 961 GCCTGCTGGGGCAAAACCAACGAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCCAGCGGTGA 1020
QY 1021 AGGCAATCAAGTGTGTCCTCTCACTGGTGAAGAAAAACCAACCTTGGCGGCCCAATA 1080
DB 1021 AGGCAATCAAGTGTGTCCTCTCACTGGTGAAGAAAAACCAACCTTGGCGGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCAATTAATGAGCTGGCAGCAGAGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCAATTAATGAGCTGGCAGCAGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTCACTCATTTAG 1200
DB 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTCACTCATTTAG 1200
QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGATGCGCGGTGACCAATGTTCTGGGG 1260
DB 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGATGCGCGGTGACCAATGTTCTGGGG 1260
QY 1261 TCAGGACGCATCGGAAGCTGTGGTATGGCTGTGAGGTCGTGAATTCATCTGATTAATTCG 1320
```

Db 1261 TCAGGAGGCCATCGGAAGCTGTGGTATGCTGTGCAGGTCGTAATCACTGCATAATTCG 1320
QY 1321 TGTGCGCTCAAGGGCGACTCCCGTTCTGGATAATGTTTTTTCGCGACATCAATACGGTT 1380
Db 1321 TGTGCGCTCAAGGGCGACTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCAATACGGTT 1380
QY 1381 CTGGCAAAATATTCTGAAATGAGCTGTTGACAAATTAATCATCGCTCGTATAATGTTGGA 1440
Db 1381 CTGGCAAAATATTCTGAAATGAGCTGTTGACAAATTAATCATCGCTCGTATAATGTTGGA 1440
QY 1441 ATTGTGAGCGGATACAAATTTACACAGGAACAGCCAGTCGGTTAGTGTTTTCAAGA 1500
Db 1441 ATTGTGAGCGGATACAAATTTACACAGGAACAGCCAGTCGGTTAGTGTTTTCAAGA 1500
QY 1501 GCACCTTCAACAAAGGACCATAGATTATGAAAACTGAAAGAGGTAAACTGCTAATCTCG 1560
Db 1501 GCACCTTCAACAAAGGACCATAGCATATGAANAATCGAAGAGGTAAACTGCTAATCTCG 1560
QY 1561 ATTAAACGGGATAAAGGCTATTAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAAACGGGATAAAGGCTATTAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACGGAATTTAAAGTCAACCGTTGAGCATCCGATCGGATTAACCTGGAAGAGANAATCCCAAGGTT 1680
Db 1621 ACGGAATTTAAAGTCAACCGTTGAGCATCCGATCGGATTAACCTGGAAGAGANAATCCCAAGGTT 1680
QY 1681 GCGCAACTGGCGATGGCCCTGACATTATCTTCTGGGACACAGCCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTATCTTCTGGGACACAGCCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTTGGCTGAAATCATCCCGGACAAAGCGTTCAGGACAAGCTGAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAAATCATCCCGGACAAAGCGTTCAGGACAAGCTGAT 1800
QY 1801 CGGTTTACTCGGATGCGGTACGTTTACAAACGCAAGCTGATGCTTACCCGATCGCTGTT 1860
Db 1801 CGGTTTACTCGGATGCGGTACGTTTACAAACGCAAGCTGATGCTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCCGAACCCGCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCCGAACCCGCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGATTAAGAACTGAAGGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGATTAAGAACTGAAGGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGACGGGGTATGCGTTCAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGACGGGGTATGCGTTCAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAAGAAACAAACACATGAATGACAGACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAAGAAACAAACACATGAATGACAGACCGATTAC 2160
QY 2161 TCCATCGCAGAAGCTGCTTTAATAAGGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAAGCTGCTTTAATAAGGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCACATCGACACAGCAAGTGAATTAATGCTGAACGGTACTGCGACCTTC 2280
Db 2221 GCATGGTCCACATCGACACAGCAAGTGAATTAATGCTGAACGGTACTGCGACCTTC 2280
QY 2281 AAGGGTCAACCATCAACACCTGCTGCTGGCGTGTGAGCGCAGGTATTAACGGCCGAGT 2340
Db 2281 AAGGGTCAACCATCAACACCTGCTGCTGGCGTGTGAGCGCAGGTATTAACGGCCGAGT 2340
QY 2341 CCGAACAAAGCTGGCAAAAGGTTCTCTGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAACAAAGCTGGCAAAAGGTTCTCTGAAAACTATCTGCTGACTGATGAAGGTCTG 2400

QY 2401 GAAGCGGTTAATAAGACAAACCCGCTGGTGCCTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCCGCTGGTGCCTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGCGGAAAGATCCACGCTATTGCGCGCACCAATGGAACGCGCCGAGAAAGGTGAATCATG 2520
Db 2461 TTGCGGAAAGATCCACGCTATTGCGCGCACCAATGGAACGCGCCGAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCGCGAGATGTCGCTTTCTGSTATGCCGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCGCGAGATGTCGCTTTCTGSTATGCCGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGTCTG 2640
Db 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGTCTG 2640
QY 2641 AACCAACAAACAATAAACAATAACAACAACCTCGGATCGAGGAAGGATTTTCAGAATTC 2700
Db 2641 AACCAACAAACAATAAACAATAACAACAACCTCGGATCGAGGAAGGATTTTCAGAATTC 2700

RESULT 13
ADO23649
ID ADO23649 standard; DNA; 7370 BP.
XX AC ADO23649;
XX AC ADO23649;
DT 01-JUL-2004 (first entry)
XX DNA encoding MBP-ToxoP30MIX5 fusion protein.
DE P30 antigen; Toxo30del3C; Toxo30del2C; ToxoP30 MIX1;
XX MBP-Toxo30del3C(52-300aa); MBP-Toxo30del4C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; gene.
XX Toxoplasma gondii.
OS Synthetic.
XX US2004067239-A1.
PN 08-APR-2004.
PD 02-OCT-2002; 2002US-00263153.
PF 02-OCT-2002; 2002US-00263153.
XX 02-OCT-2002; 2002US-00263153.
PR (MAIN/) MAINE G T.
XX (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
XX (BLIE/) BLIESE T R.
PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX WPI; 2004-304563/28.
DR P-PSDB; ADO23650.
XX Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del3C,
PT Toxo30del2C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
PT antibodies to Toxoplasma gondii.
XX Example 5; Fig 35; 114pp; English.
XX The invention relates to a purified P30 antigen (I) chosen from 3 fully
CC defined Toxo30del3C, Toxo30del2C and ToxoP30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxo30del2C P30 antigen sequence, or comprising an amino acid sequence of
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxo30del3C P30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxo30del3C(52-300aa), MBP-Toxo30del4C(52-294aa) and MBP-Toxo30MIX1
CC fusion proteins. (I) is useful for detecting the presence of IgM

CC antibodies to *Toxoplasma gondii* in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (i) and detecting the presence of (i)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC Toxop30 fusion protein of the invention.
XX
SQ Sequence 7370 BP; 1900 A; 1857 C; 1950 G; 1663 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY	1	CGGACACCATCAAGTGGCGAAGCTTTTCGGGTATGCGATGATAGCGCCGGAGAGA	60
DB	1	CGGACACCATCAAGTGGCGAAGCTTTTCGGGTATGCGATGATAGCGCCGGAGAGA	60
QY	61	GTCNAATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGCGCAGAGTATGCG	120
DB	61	GTCNAATTCAGGGTGGTGAATGTGAACACAGTAACTTATACGATGCGCAGAGTATGCG	120
QY	121	GTGTCCTTATCAGACCGTTTCGGGTGGTGAACACAGGCCAGCCAGCTTTCTGCGAATA	180
DB	121	GTGTCCTTATCAGACCGTTTCGGGTGGTGAACACAGGCCAGCCAGCTTTCTGCGAATA	180
QY	181	CGCGGAAAGTGAAGCGGCGATGGCGAGCTGAATTAACATTCCTCCAAACCGCGTGGC	240
DB	181	CGCGGAAAGTGAAGCGGCGATGGCGAGCTGAATTAACATTCCTCCAAACCGCGTGGC	240
QY	241	AACAACTCGCGGCAACAGTCTGCTGATTTGGGCTTGCACCTCCAGTCTGGCCCTGC	300
DB	241	AACAACTCGCGGCAACAGTCTGCTGATTTGGGCTTGCACCTCCAGTCTGGCCCTGC	300
QY	301	ACGCGCGTGCAGAAATGTTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAG	360
DB	301	ACGCGCGTGCAGAAATGTTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAG	360
QY	361	TGTTGCTCGATGTAGAACGAGCGGCTGAAAGCTTGAAGCGGCTGCAATTC	420
DB	361	TGTTGCTCGATGTAGAACGAGCGGCTGAAAGCTTGAAGCGGCTGCAATTC	420
QY	421	TTCTCGCGCAACGCTGAGTGGCTGATCATTAATCTATCGCTGATGACCAAGGATGCA	480
DB	421	TTCTCGCGCAACGCTGAGTGGCTGATCATTAATCTATCGCTGATGACCAAGGATGCA	480
QY	481	TTGCTGTGGAAGCTCCCTGCAATAATGTTCGGCGTATTTCTTGATGTCTCTGACCA	540
DB	481	TTGCTGTGGAAGCTCCCTGCAATAATGTTCGGCGTATTTCTTGATGTCTCTGACCA	540
QY	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGGGCTGGGAGCATC	600
DB	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGGGCTGGGAGCATC	600
QY	601	TGTTGCGATTTGGTCCAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTCTGCTCG	660
DB	601	TGTTGCGATTTGGTCCAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTCTGCTCG	660
QY	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGAT	720
DB	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGAT	720
QY	721	CGGAACGGGAGCGGACCTGGAGTGCATGTCCGGTTTTCACAAACCATGCAATGCTGA	780
DB	721	CGGAACGGGAGCGGACCTGGAGTGCATGTCCGGTTTTCACAAACCATGCAATGCTGA	780
QY	781	ATGAGGGATCGTCCCACTGCGATGCTGGTTGGTCCACAGATCGGCTGGCGGCA	840
DB	781	ATGAGGGATCGTCCCACTGCGATGCTGGTTGGTCCACAGATCGGCTGGCGGCA	840
QY	841	TGCGCGCATTAACAGATCGCGGCTGGTGGTGGGATATCTCGGTAGTGGGATACG	900
DB	841	TGCGCGCATTAACAGATCGCGGCTGGTGGTGGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAAGACAGCTCATGTTTATATCCCGCGCTTAACCAACATCAACAGGATTTTC	960

DB	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAACAGGATTTTC	960
QY	961	GCCTGCTGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCAGGCGGTGA	1020
DB	961	GCCTGCTGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCAGGCGGTGA	1020
QY	1021	AGGGCAATCAGCTGTGTCCTCTCACTGCTGTAAGAAAGAAACCAACCTGGCGCCAATA	1080
DB	1021	AGGGCAATCAGCTGTGTCCTCTCACTGCTGTAAGAAAGAAACCAACCTGGCGCCAATA	1080
QY	1081	CGCAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGCAGCTGCACGACGAGTTT	1140
DB	1081	CGCAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAATGCAGCTGCACGACGAGTTT	1140
QY	1141	CCGACTGGAAGGCGGCGAGTGCAGCGCAACCGCAATTAATGTGAGTTAGCTCACTCAT	1200
DB	1141	CCGACTGGAAGGCGGCGAGTGCAGCGCAACCGCAATTAATGTGAGTTAGCTCACTCAT	1200
QY	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGCGTGCACCAATCTCTG	1260
DB	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGCGTGCACCAATCTCTG	1260
QY	1261	TCAGGACGCAATCGGAAGCTGTGATGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT	1320
DB	1261	TCAGGACGCAATCGGAAGCTGTGATGCTGTGAGCTGTGAGCTGTGAGCTGTGAGCT	1320
QY	1321	TGTGCTCAAGCGCACTCCCGTCTTGATTAATGTTTTCGCGCGACATCATACCGTT	1380
DB	1321	TGTGCTCAAGCGCACTCCCGTCTTGATTAATGTTTTCGCGCGACATCATACCGTT	1380
QY	1381	CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATCTG	1440
DB	1381	CTGGCAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATCTG	1440
QY	1441	ATTGTGAGCGATTAACAAATTTTTCACAGGAAACGCGCTGCTGAGTGTGTTTTCAC	1500
DB	1441	ATTGTGAGCGATTAACAAATTTTTCACAGGAAACGCGCTGCTGAGTGTGTTTTCAC	1500
QY	1501	GCATTTTCAACCAAGGACCATAGATTAATGAAAGCTGAAAGGTAAGCTGTAATCTGG	1560
DB	1501	GCATTTTCAACCAAGGACCATAGATTAATGAAAGCTGAAAGGTAAGCTGTAATCTGG	1560
QY	1561	ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAACTCGGTGCGTAAGAAATTC	1620
DB	1561	ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAACTCGGTGCGTAAGAAATTC	1620
QY	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAG	1680
DB	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAG	1680
QY	1681	CGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACCGCTTTGGTGGCTAC	1740
DB	1681	CGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCTGTTGGCTGAAATCAACCGGCAAAAGCGTTCCAGGACAAGCTGAT	1800
DB	1741	GCTCAATCTGGCTGTTGGCTGAAATCAACCGGCAAAAGCGTTCCAGGACAAGCTGAT	1800
QY	1801	CGTTTACCTGGGATGCGGTACGTTTAAACCGCAAGCTGATTTGCTTACCGATCGCT	1860
DB	1801	CGTTTACCTGGGATGCGGTACGTTTAAACCGCAAGCTGATTTGCTTACCGATCGCT	1860
QY	1861	GAAGGTTATCGCTGATTTTAAACAAAGATCTGCTGCGAAACCGCCCAAAACCTGG	1920
DB	1861	GAAGGTTATCGCTGATTTTAAACAAAGATCTGCTGCGAAACCGCCCAAAACCTGG	1920
QY	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGAGTAAAGAGCGCTGATGTTCA	1980
DB	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGAGTAAAGAGCGCTGATGTTCA	1980
QY	1981	CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGCTGACGCGGGTATGCTGTTCA	2040

QY 541 CACCCATCAACAGTATTATTTCTCCATGAAGACGGTACCGGACTGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCATGAAGACGGTACCGGACTGGCGTGGAGCATC 600
QY 601 TGGTCGCATGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db 601 TGGTCGCATGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTGGCTGGCTGGCGTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGGCTGGCTGGCGTGGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGACCGGAGGACGCTGGAGTGCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAAACGGGAAGGACGCTGGAGTGCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTGGCCACGATCAGATGGCGCTGGCGGCA 840
Db 781 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTGGCCACGATCAGATGGCGCTGGCGGCA 840
QY 841 TGGCGGCCATTAACGAGTCCGGGCTGGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
Db 841 TGGCGGCCATTAACGAGTCCGGGCTGGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACGAGTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAGAAACCCACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAGAAACCCACCTGGCGCCCAATA 1080
QY 1081 CGCAAAACGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGCTGGCAGCAGAGTTT 1140
Db 1081 CGCAAAACGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGCTGGCAGCAGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGGCAGTGCAGCGCAACGAATTAATGTGAGTTAGTCACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGCAGCGCAACGAATTAATGTGAGTTAGTCACTCATTAG 1200
QY 1201 GCACAATTCATGTTTGAAGCTTTATCATCGACTGCAGGTCGACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTCATGTTTGAAGCTTTATCATCGACTGCAGGTCGACCAATGCTTCTGGCG 1260
QY 1261 TCAGGCGGCATCGGAAGCTGGTATGGCTGTGAGGTCGTAATCACTGCATTAATTCG 1320
Db 1261 TCAGGCGGCATCGGAAGCTGGTATGGCTGTGAGGTCGTAATCACTGCATTAATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGTCTGGATATGTTTTTGGCGCGACATCAACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTGGATATGTTTTTGGCGCGACATCAACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
QY 1441 ATTGTAGCGGATACAAATTCACACAGAAACAGCGAGTCGGTTAGGTGTTTTACGA 1500
Db 1441 ATTGTAGCGGATACAAATTCACACAGAAACAGCGAGTCGGTTAGGTGTTTTACGA 1500
QY 1501 GCACCTTCCACCAACAGGACCATAGCATATGAAAATCGAAGAGGTAACCTGTAATCTGG 1560
Db 1501 GCACCTTCCACCAACAGGACCATAGCATATGAAAATCGAAGAGGTAACCTGTAATCTGG 1560
QY 1561 ATTTAAGCGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTTAAGCGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT 1680

Db 1621 ACCGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT 1680
QY 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTGTGGCTGAATTAACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTGTGGCTGAATTAACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGGCTAGCTTACAACGCGAAGCTGATTGCTTACCCGATCCTGTT 1860
Db 1801 CCGTTTACCTGGGATGGCTAGCTTACAACGCGAAGCTGATTGCTTACCCGATCCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAGCGAAAGSTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAGCGAAAGSTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCCACCTGGCGCTGATTGCTGTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCCACCTGGCGCTGATTGCTGTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGGCAAGTACGACATTAAGACCGTGGCGCTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGGCAAGTACGACATTAAGACCGTGGCGCTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACATGAAATGCGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACATGAAATGCGAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCCCTTTAATAAGCGGCAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCCCTTTAATAAGCGGCAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGCTCAACATCGACACCAAGTGAATTAATGCTGTAACCGGTACTGCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACCAAGTGAATTAATGCTGTAACCGGTACTGCGGACCTTC 2280
QY 2281 AAGGTCACCAATCCAAACCGTTGCTGGCGTCTGAGCGCAGGTATTAACCGCCCACT 2340
Db 2281 AAGGTCACCAATCCAAACCGTTGCTGGCGTCTGAGCGCAGGTATTAACCGCCCACT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCCGCTGGGTGGCTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCCGCTGGGTGGCTAGCGCTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCACGATTATTCGCCCAACCATGGAAGAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTATTCGCCCAACCATGGAAGAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CGGAACATCCCGCAGATGTCGGCTTCTGGTATGCGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CGGAACATCCCGCAGATGTCGGCTTCTGGTATGCGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACCAACAACAATAACAATAACAACCTTCGGGATCGAGGGAAGGATTTCAAGATTC 2700
Db 2641 AACCAACAACAATAACAATAACAACCTTCGGGATCGAGGGAAGGATTTCAAGATTC 2700

ADO23594
 ID ADO23594 standard; DNA; 7403 BP.
 XX AC ADO23594;
 XX DT 01-JUL-2004 (first entry)
 XX DE DNA encoding MBP-Toxop30del2(52-311aa) fusion protein.
 XX KW P30 antigen; Toxo30del3C; Toxo30del2C; ToxoP30 MIX1;
 KW MBP-Toxo30del3C(52-300aa); MBP-Toxo30del4C(52-294aa); MBP-Toxo30MIX1;
 KW Toxoplasma gondii; ds, gene.
 XX OS Toxoplasma gondii.
 OS Synthetic.
 XX PN US2004067239-A1.
 XX PD 08-APR-2004.
 XX PF 02-OCT-2002; 2002US-00263153.
 XX PR 02-OCT-2002; 2002US-00263153.
 XX PA (MAIN/) MAINE G T.
 PA (PATE/) PATEL C B.
 PA (GINS/) GINSBURG S R.
 PA (BLIE/) BLIESE T R.
 PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
 XX WPI: 2004-304563/28.
 DR P-PSDB; ADO23595.
 XX
 PT Novel purified polypeptide having sequence identity to amino acid
 sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del3C,
 PT Toxo30del2C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
 PT antibodies to Toxoplasma gondii.
 XX
 PS Example 2; Fig 11; 114pp; English.
 XX
 CC The invention relates to a purified P30 antigen (I) chosen from 3 fully
 CC defined Toxo30del3C, Toxo30del2C and ToxoP30 MIX1 sequences, having 1-6
 CC amino acids added to the C-terminus of the amino acid sequence of
 CC Toxo30del2C P30 antigen sequence, or comprising an amino acid sequence in
 CC which at least one of the five C-terminal cysteine amino acids of the
 CC amino acid sequence of Toxo30del3C P30 antigen sequence is substituted
 CC with alanine, or comprising the amino acid sequence chosen from MBP-
 CC Toxo30del3C(52-300aa), MBP-Toxo30del4C(52-294aa) and MBP-Toxo30MIX1
 CC fusion proteins. (I) is useful for detecting the presence of IgM
 CC antibodies to Toxoplasma gondii in a test sample, which involves
 CC contacting the test sample suspected of containing the IgM antibodies
 CC with a composition comprising (I) and detecting the presence of (I)/IgM
 CC antibody complexes. The present sequence represents DNA encoding a MBP-
 CC ToxoP30 fusion protein of the invention.
 XX
 SQ Sequence 7403 BP; 1907 A; 1863 C; 1961 G; 1672 T; 0 U; 0 Other;
 Query Match 81.5%; Score 2690.4; DB 12; Length 7403;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CCGACACCATCGAATGGTGGCAAAACCTTTCGGGTATGGCATGATAGGCCCGCGGAAGAGA 60
 DB 1 CCGACACCATCGAATGGTGGCAAAACCTTTCGGGTATGGCATGATAGGCCCGCGGAAGAGA 60
 QY 61 GTCAATTGAGGGTGGTGAATGTGAACACGTAACGTTATACGATGTGCGAGAGTATCCCG 120
 DB 61 GTCAATTGAGGGTGGTGAATGTGAACACGTAACGTTATACGATGTGCGAGAGTATCCCG 120
 QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGGAAACCGAGCCAGCCAGCGTTCTGCGAAAA 180
 DB 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGGAAACCGAGCCAGCCAGCGTTCTGCGAAAA 180

QY 181 CGCGGAAAAAGTGAAGCGCGGATGCGGAGCTGAATTACATTCCTCAACCGCGTGGCAC 240
 DB 181 CGCGGAAAAAGTGAAGCGCGGATGCGGAGCTGAATTACATTCCTCAACCGCGTGGCAC 240
 QY 241 AACAACTGCGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGCCCTGC 300
 DB 241 AACAACTGCGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGCCCTGC 300
 QY 301 AGCGCGGTCCGCAAAATGTGTCGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
 DB 301 AGCGCGGTCCGCAAAATGTGTCGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
 QY 361 TGGTGGTGTGATGGTAGAAGCGCGCTGGAAGCTGTAAAGCGCGGTGCACAATC 420
 DB 361 TGGTGGTGTGATGGTAGAAGCGCGCTGGAAGCTGTAAAGCGCGGTGCACAATC 420
 QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCACTAATTAATCTCGCTGATGACACGATGCCA 480
 DB 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCACTAATTAATCTCGCTGATGACACGATGCCA 480
 QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
 DB 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
 QY 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 600
 DB 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 600
 QY 601 TGGTCCGATTGGGTACCAGCAAAATCGCGTGTGTAGCGGGCCCATTAAGTTCTGTCGCG 660
 DB 601 TGGTCCGATTGGGTACCAGCAAAATCGCGTGTGTAGCGGGCCCATTAAGTTCTGTCGCG 660
 QY 661 CGCGTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATCAGCCGATAG 720
 DB 661 CGCGTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATCAGCCGATAG 720
 QY 721 CGGAAACGGGAAGCGACTGGAGTGCCATGTCGGGTTTCAACAAACCATGCAATGCTGA 780
 DB 721 CGGAAACGGGAAGCGACTGGAGTGCCATGTCGGGTTTCAACAAACCATGCAATGCTGA 780
 QY 781 ATGAGGCAATGTTTCCACTGCGATGCTGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
 DB 781 ATGAGGCAATGTTTCCACTGCGATGCTGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
 QY 841 TGGCGCCATTACCGAGTCCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTAG 900
 DB 841 TGGCGCCATTACCGAGTCCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTAG 900
 QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
 DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
 QY 961 GCCTGTGGGGCAAAACAGGTGGACCGGTGTCGCAACTCTCTCAGGGCCAGCGGTGA 1020
 DB 961 GCCTGTGGGGCAAAACAGGTGGACCGGTGTCGCAACTCTCTCAGGGCCAGCGGTGA 1020
 QY 1021 AGGGCAATCAGCTGTGCGCGTCTCACTGGTGAAGAAACCAACCCCTGGGCGCCCAATA 1080
 DB 1021 AGGGCAATCAGCTGTGCGCGTCTCACTGGTGAAGAAACCAACCCCTGGGCGCCCAATA 1080
 QY 1081 CGCAAAACCGCTCTCCCGCGGTGGCGGATTTCAATTAATGCGAGTGGGACGACAGGTTT 1140
 DB 1081 CGCAAAACCGCTCTCCCGCGGTGGCGGATTTCAATTAATGCGAGTGGGACGACAGGTTT 1140
 QY 1141 CCGGCTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTGAGTTAGTCTACTATTAG 1200
 DB 1141 CCGGCTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTGAGTTAGTCTACTATTAG 1200
 QY 1201 GCACAAATTCATGTTTGAAGCTTATCATCGACTGCAAGCGTGCACCAATGCTTCTGGCG 1260
 DB 1201 GCACAAATTCATGTTTGAAGCTTATCATCGACTGCAAGCGTGCACCAATGCTTCTGGCG 1260

Qy 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTGTAATCACTGCATTAATTCG 1320
Db 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTGTAATCACTGCATTAATTCG 1320
Qy 1321 TGTGCTCAAGCGGCACTCCCGTCTCGGATTAATGTTTTTGGCCGACATCATAAACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTCTCGGATTAATGTTTTTGGCCGACATCATAAACGGTT 1380
Qy 1381 CTGGCAAAATATCTGAAATGAGCTGTTCACAAATTAATCATCGGCTCGTATATGTTGTGA 1440
Db 1381 CTGGCAAAATATCTGAAATGAGCTGTTCACAAATTAATCATCGGCTCGTATATGTTGTGA 1440
Qy 1441 ATTGTGAGCGGATAACAATTTACACAGGAACACCCAGTCGGTTAGTGTTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAATTTACACAGGAACACCCAGTCGGTTAGTGTTTTTCACGA 1500
Qy 1501 GCACCTTACCAACAGGACCATAGATTAATGAAACTGGAAGAGTAAACTGTTATCTGG 1560
Db 1501 GCACCTTACCAACAGGACCATAGATTAATGAAACTGGAAGAGTAAACTGTTATCTGG 1560
Qy 1561 ATTTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAAGTCGGTAAGAAAATTCGAGAAAGAT 1620
Db 1561 ATTTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAAGTCGGTAAGAAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCGGATTAACCTGGAAGAGAAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCGGATTAACCTGGAAGAGAAAATTCACAGGTT 1680
Qy 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTCGGGCACACAGCCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTCGGGCACACAGCCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTCCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTCCTTACCCGATCGCTGT 1860
Qy 1861 GAAGGTTTATCGCTGATTTATAACAAAGATCTGCTGCCGACACCCGCCAAACCTGGGAA 1920
Db 1861 GAAGGTTTATCGCTGATTTATAACAAAGATCTGCTGCCGACACCCGCCAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCGCTCGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTCGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGGCGGCTGATGCTGAGCGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGGCTGATGCTGAGCGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAGACGCAAGTACGACATTAAGACGTTGGCGGTGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAGACGCAAGTACGACATTAAGACGTTGGCGGTGATTAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGTTGACCTGATTAAGAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAAGAAACAAACACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAAAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAAAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCCACATCGACACGCAAGTGAATTAATGTTAAACGCTGCGGACCTTC 2280
Db 2221 GCATGGTCCACATCGACACGCAAGTGAATTAATGTTAAACGCTGCGGACCTTC 2280
Qy 2281 AAGGTTCAACCATCCAAACCGTTGTTGGCGTGTGAGCGCAGGTTAATAACGCCGCCAGT 2340
Db 2281 AAGGTTCAACCATCCAAACCGTTGTTGGCGTGTGAGCGCAGGTTAATAACGCCGCCAGT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400

Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCTGAAAACTATCTGCTGACTGATGAAGGTCG 2400
Qy 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTCCCGTAGCGCTGAAAGTCTTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTCCCGTAGCGCTGAAAGTCTTTACGAGGAAGAG 2460
Qy 2461 TTGGCGAAAAGATCCACGTTATTTGCCGCCACCATGGAACAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAAGATCCACGTTATTTGCCGCCACCATGGAACAAACGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTCTTGTTATGCGGTGCTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTTGTTATGCGGTGCTACTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG 2640
Qy 2641 AACAAACAAACAATTAACAATAACAAACCTCGGATCGAGGGAAGGATTTTCAAAATTC 2700
Db 2641 AACAAACAAACAATTAACAATAACAAACCTCGGATCGAGGGAAGGATTTTCAAAATTC 2700

Search completed: August 1, 2005, 04:28:30
Job time : 1104.6 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 352.4 Seconds
(without alignments)
15322.682 Million cell updates/sec

Title: US-09-765-555B-14
Perfect score: 3300
Sequence: 1 ccgacaccatcgatgtgc.....acgacgttcgcgactaagct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2700	81.8	7475	US-08-971-036-1	Sequence 1, Appli
2	2700	81.8	7475	US-09-096-570-1	Sequence 1, Appli
3	2700	81.8	7475	US-09-263-617B-1	Sequence 1, Appli
c 4	1417.6	43.0	5926	US-09-027-169-3	Sequence 3, Appli
5	1362	41.3	5201	US-09-640-882-2	Sequence 2, Appli
6	1362	41.3	5201	US-09-640-882-3	Sequence 3, Appli
7	1204	36.5	4557	US-08-778-717-5	Sequence 5, Appli
8	1200.8	36.4	2832	US-08-148-675A-2	Sequence 2, Appli
c 9	1199.2	36.3	5248	US-08-487-281A-18	Sequence 18, Appli
c 10	1199.2	36.3	5248	PT-US96-05611A-21	Sequence 21, Appli
c 11	1199.2	36.3	5312	US-10-263-103-35	Sequence 35, Appli
c 12	1199.2	36.3	5443	US-08-929-967-1	Sequence 1, Appli
c 13	1199.2	36.3	5502	US-09-702-705-785	Sequence 785, App
c 14	1199.2	36.3	5502	US-09-736-457-785	Sequence 785, App
c 15	1199.2	36.3	5502	US-09-614-124B-785	Sequence 785, App
c 16	1199.2	36.3	5502	US-09-671-325-785	Sequence 785, App
c 17	1199.2	36.3	5502	US-09-589-184-785	Sequence 785, App
c 18	1199.2	36.3	5502	US-09-658-824-785	Sequence 785, App
c 19	1199.2	36.3	5616	US-08-929-967-3	Sequence 3, Appli
c 20	1199.2	36.3	5873	US-09-695-437A-62	Sequence 62, Appli
c 21	1199.2	36.3	6353	US-09-702-705-784	Sequence 784, App
c 22	1199.2	36.3	6353	US-09-736-457-784	Sequence 784, App
c 23	1199.2	36.3	6353	US-09-614-124B-784	Sequence 784, App
c 24	1199.2	36.3	6353	US-09-671-325-784	Sequence 784, App
c 25	1199.2	36.3	6353	US-09-589-184-784	Sequence 784, App
c 26	1199.2	36.3	6353	US-09-658-824-784	Sequence 784, App
27	1199.2	36.3	6361	US-08-646-538-7	Sequence 7, Appli

c 28	1199.2	36.3	6361	3	US-09-503-222-7	Sequence 7, Appli
c 29	1199.2	36.3	6727	2	US-08-125-462-2	Sequence 2, Appli
c 30	1199.2	36.3	6727	2	US-08-891-848-2	Sequence 2, Appli
c 31	1199.2	36.3	6799	2	US-08-125-462-5	Sequence 5, Appli
c 32	1199.2	36.3	6799	2	US-08-891-848-5	Sequence 5, Appli
c 33	1199.2	36.3	7659	3	US-09-128-314-4	Sequence 4, Appli
c 34	1199.2	36.3	7676	3	US-09-056-556-213	Sequence 213, App
c 35	1199.2	36.3	7676	3	US-09-072-596-208	Sequence 208, App
c 36	1199.2	36.3	7676	4	US-09-072-967-213	Sequence 213, App
c 37	1199.2	36.3	7676	4	US-09-287-849-9	Sequence 9, Appli
c 38	1199.2	36.3	8031	3	US-09-643-597-254	Sequence 254, App
c 39	1199.2	36.3	8031	4	US-09-480-884A-254	Sequence 254, App
c 40	1199.2	36.3	8031	4	US-09-542-615A-254	Sequence 254, App
c 41	1199.2	36.3	8031	4	US-09-606-421B-254	Sequence 254, App
c 42	1199.2	36.3	8031	4	US-09-476-496A-254	Sequence 254, App
c 43	1199.2	36.3	8031	4	US-09-630-940B-254	Sequence 254, App
c 44	1199.2	36.3	8157	3	US-09-128-314-3	Sequence 3, Appli
c 45	1199.2	36.3	8501	3	US-08-793-900-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-971-036-1
; Sequence 1, Application US/08971036
; Patent No. 5866684
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R
; APPLICANT: Hurst, David N
; APPLICANT: Jones, Philip S
; APPLICANT: Kay, Paul B
; APPLICANT: Raynham, Tony M
; APPLICANT: Wilson, Francis X
; TITLE OF INVENTION: Amino Acid Derivatives
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: N.J.
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/971.036
; FILING DATE: 14-NOV-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9623908.2
; FILING DATE: 18-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kreisler, Lewis J
; REGISTRATION NUMBER: 38522
; REFERENCE/DOCKET NUMBER: RAN 4430/073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (973) 235-4387
; TELEFAX: (973) 235-2363
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
US-08-971-036-1

Query Match 81.8% ; Score 2700 ; DB 2 ; Length 7475 ;

Best Local Similarity 100.0%; Pred. No. 0;			
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	CGGACACCATCAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGAGAGA	60
Db	1	CGGACACCATCAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGAGAGA	60
QY	61	GTCAATTGAGGTGGTGAATGTGAACCAAGTAACTGATGTCGACAGTATGCGG	120
Db	61	GTCAATTGAGGTGGTGAATGTGAACCAAGTAACTGATGTCGACAGTATGCGG	120
QY	121	GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGGCGAGCACTGTTCTCGGAAA	180
Db	121	GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGGCGAGCACTGTTCTCGGAAA	180
QY	181	CGCGGAAAAAGTGAAGCGGATGCGGAGCTGAATTAATTAATCCCAACCGGTGCAC	240
Db	181	CGCGGAAAAAGTGAAGCGGATGCGGAGCTGAATTAATTAATCCCAACCGGTGCAC	240
QY	241	AACAACTGCGCGGCAACAGTCTGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
Db	241	AACAACTGCGCGGCAACAGTCTGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
QY	301	ACGCGCGTGCAGAAATGTCGCGGCGATTAATCTCGCGCGATCAACTGGGTGCCAGC	360
Db	301	ACGCGCGTGCAGAAATGTCGCGGCGATTAATCTCGCGCGATCAACTGGGTGCCAGC	360
QY	361	TGTTGGTTCGATGTAAGCAAGCGCGCTGGAAGCCTGTAAGCGCGGTGCACAATC	420
Db	361	TGTTGGTTCGATGTAAGCAAGCGCGCTGGAAGCCTGTAAGCGCGGTGCACAATC	420
QY	421	TTCTGCGCAACCGGTCACTAGTGGGTGATCAATTAATCTATCCGCTGGATGACCAAGTGC	480
Db	421	TTCTGCGCAACCGGTCACTAGTGGGTGATCAATTAATCTATCCGCTGGATGACCAAGTGC	480
QY	481	TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGGTTATTTCTTGATGTTCTGACGAGA	540
Db	481	TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGGTTATTTCTTGATGTTCTGACGAGA	540
QY	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGCACTGGCGGTGAGCATC	600
Db	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGCACTGGCGGTGAGCATC	600
QY	601	TGTTGCGAATGGGTCAACAGCAATTCGCGTGTGTAGCGGGCCCAATTAAGTTCTGTCGG	660
Db	601	TGTTGCGAATGGGTCAACAGCAATTCGCGTGTGTAGCGGGCCCAATTAAGTTCTGTCGG	660
QY	661	CGCGTCTGCGTCTGCGTGGCATAAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Db	661	CGCGTCTGCGTCTGCGTGGCATAAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAACCGGAAGGCGACTGGAGTGCCATGTCGGGTTTTCACAAACCATGCAAAATGCTGA	780
Db	721	CGGAACCGGAAGGCGACTGGAGTGCCATGTCGGGTTTTCACAAACCATGCAAAATGCTGA	780
QY	781	ATGAGGGATCGTTCCCATCGCATGCTGTTGTCGCAACGATCAGATCGCGCTGGCGCAA	840
Db	781	ATGAGGGATCGTTCCCATCGCATGCTGTTGTCGCAACGATCAGATCGCGCTGGCGCAA	840
QY	841	TGCGCGCATTTACCGAGTTCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGCGCATTTACCGAGTTCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC	960
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC	960
QY	961	GCCTGCTGGGCAACACAGGTGACCGCTTCTGCAACTCTCTCAGGGCGACCGGTGA	1020
Db	961	GCCTGCTGGGCAACACAGGTGACCGCTTCTGCAACTCTCTCAGGGCGACCGGTGA	1020
QY	1021	AGGCAATACAGCTGTTGCCGCTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA	1080
Db	1021	AGGCAATACAGCTGTTGCCGCTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA	1080

Db	1021	AGGCAATACAGCTGTTGCCGCTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA	1080
QY	1081	CGAAAAACCGCTCTCCCGCGGCTTGGCGGATTCATTTAATGCACTGGCAAGAGAGT	1140
Db	1081	CGAAAAACCGCTCTCCCGCGGCTTGGCGGATTCATTTAATGCACTGGCAAGAGAGT	1140
QY	1141	CCGCACTGGAAGAGGCGAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTCATTTAG	1200
Db	1141	CCGCACTGGAAGAGGCGAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTCATTTAG	1200
QY	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAACGCTGACCAATGCTTCTGGCG	1260
Db	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAACGCTGACCAATGCTTCTGGCG	1260
QY	1261	TCAGGAGCCATCGGAAGCTTGGTATGGCTGTGACGCTGCTAAATCACTGCAATAATTCG	1320
Db	1261	TCAGGAGCCATCGGAAGCTTGGTATGGCTGTGACGCTGCTAAATCACTGCAATAATTCG	1320
QY	1321	TGTCGCTCAAGGCGCACTCCCGTTCGATAATGTTTTTGGCGCGACATCAATAACGGTT	1380
Db	1321	TGTCGCTCAAGGCGCACTCCCGTTCGATAATGTTTTTGGCGCGACATCAATAACGGTT	1380
QY	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGGA	1440
Db	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGGA	1440
QY	1441	ATTGTGAGCGGATAACAAATTTCAACAGGAAAAACGCGAGTCCGTTTATGTTTTCACGA	1500
Db	1441	ATTGTGAGCGGATAACAAATTTCAACAGGAAAAACGCGAGTCCGTTTATGTTTTCACGA	1500
QY	1501	GCATCTCACCAACAAAGGACCATAGATTATGAAACTGGAAGAGTAACTGGTAACTCGG	1560
Db	1501	GCATCTCACCAACAAAGGACCATAGATTATGAAACTGGAAGAGTAACTGGTAACTCGG	1560
QY	1561	ATTAAACCGGCGATAAAGGCTATAACCGTCTCGTGAAGTCCGTTAAGAAATTCGAGAA	1620
Db	1561	ATTAAACCGGCGATAAAGGCTATAACCGTCTCGTGAAGTCCGTTAAGAAATTCGAGAA	1620
QY	1621	ACCGGAATTAAGTCAACGTTGAGCATCCGGATAAACTGGAAGAGTAACTCCCAAGTTT	1680
Db	1621	ACCGGAATTAAGTCAACGTTGAGCATCCGGATAAACTGGAAGAGTAACTCCCAAGTTT	1680
QY	1681	CGCGCAACTGGCGATGGCCCTGACATTTCTCTGGGCAACGACCGCTTGGTGGCTAC	1740
Db	1681	CGCGCAACTGGCGATGGCCCTGACATTTCTCTGGGCAACGACCGCTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCTGTTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAGCTGTAT	1800
Db	1741	GCTCAATCTGGCTGTTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAGCTGTAT	1800
QY	1801	CCGTTTACCTGGGATGCGGTACGTTTACACGCAAGCTGATTCCTTACCCGATCGCTGTT	1860
Db	1801	CCGTTTACCTGGGATGCGGTACGTTTACACGCAAGCTGATTCCTTACCCGATCGCTGTT	1860
QY	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTCTCGCGAACCCCGCAAAACCTGGAA	1920
Db	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTCTCGCGAACCCCGCAAAACCTGGAA	1920
QY	1921	GAGATCCCGCGCTGATTAAGAACTGAAAGCGGAAAGGTAAAGCGGCTGTATGTTCAAC	1980
Db	1921	GAGATCCCGCGCTGATTAAGAACTGAAAGCGGAAAGGTAAAGCGGCTGTATGTTCAAC	1980
QY	1981	CTCAAGAACCGGTACTTCACTGGCGCTGATTTGCTGACGCGGCTTATGTTTCAAG	2040
Db	1981	CTCAAGAACCGGTACTTCACTGGCGCTGATTTGCTGACGCGGCTTATGTTTCAAG	2040
QY	2041	TATGAAAAACGCAAGTACGACATTAAGACCTGGGCTGGATAAACGCTGGCGGAAAGCG	2100
Db	2041	TATGAAAAACGCAAGTACGACATTAAGACCTGGGCTGGATAAACGCTGGCGGAAAGCG	2100
QY	2101	GGTCTGACCTTCTGCTGCTGCTGATTAAGAAACCAACACATGAATGAGACCGATTAC	2160
Db	2101	GGTCTGACCTTCTGCTGCTGCTGATTAAGAAACCAACACATGAATGAGACCGATTAC	2160

QY 841 TCGCGGCCATTACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGTAGTGGGATACG 900
DB 841 TCGCGGCCATTACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCCATCAAAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCCATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACGAGTGGGACCGCTTCTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
DB 961 GCCTGCTGGGCAACACGAGTGGGACCGCTTCTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGGGAAAGAAACCAACCTGGCGGCCAATA 1080
DB 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGGGAAAGAAACCAACCTGGCGGCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGACGTGGCACACAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGACGTGGCACACAGGTTT 1140
QY 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACCAATTAATGAGTTAGTCACTCACTTAG 1200
DB 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACCAATTAATGAGTTAGTCACTCACTTAG 1200
QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGACGCTGCACCAATGCTTCTGGCG 1260
DB 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGACGCTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACGCAATCGAAAGCTGTGATGCTGTGAGCTGCAAGCTGCAATCACTGATTAATTCG 1320
DB 1261 TCAGGACGCAATCGAAAGCTGTGATGCTGTGAGCTGCAAGCTGCAATCACTGATTAATTCG 1320
QY 1321 TGTGCTCTCAAGCGGCACTCCCGTCTCGGATTAATGTTTTCGCGGACATCAACCGTT 1380
DB 1321 TGTGCTCTCAAGCGGCACTCCCGTCTCGGATTAATGTTTTCGCGGACATCAACCGTT 1380
QY 1381 CTGGCAAAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
DB 1381 CTGGCAAAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
QY 1441 ATTGTGAGCGGATACAAATTTTACAGGAAACGCGAGTCCGTTAGGTGTTTTCAGGA 1500
DB 1441 ATTGTGAGCGGATACAAATTTTACAGGAAACGCGAGTCCGTTAGGTGTTTTCAGGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAAACTGAAAGCTGAACTGTAATCTGG 1560
DB 1501 GCACCTTCAACCAAGGACCATAGATTATGAAACTGAAAGCTGAACTGTAATCTGG 1560
QY 1561 ATTAACGCGATTAAGGCTATAACGCTCTGCTGAACTGAACTGAACTGAACTGAACTGAA 1620
DB 1561 ATTAACGCGATTAAGGCTATAACGCTCTGCTGAACTGAACTGAACTGAACTGAACTGAA 1620
QY 1621 ACCGGAATTAAGTCAACGCTTGGATCCGATTAACCTGGAAGAGAAATCCCAAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACGCTTGGATCCGATTAACCTGGAAGAGAAATCCCAAGGTT 1680
QY 1681 CGCGCAACTGCGGATGCGCTGACATTTCTTCTGGGACACGACCGCTTGGTGGCTAC 1740
DB 1681 CGCGCAACTGCGGATGCGCTGACATTTCTTCTGGGACACGACCGCTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTTGGCTGAAATCAACCGGCAACGCTTCCAGGCAAGCTGTAT 1800
DB 1741 GCTCAATCTGCGCTTGGCTGAAATCAACCGGCAACGCTTCCAGGCAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGCTGATTAACCGGCAACGCTTCCAGGCAAGCTGTAT 1860
DB 1801 CCGTTTACCTGGGATGCGCTGATTAACCGGCAACGCTTCCAGGCAAGCTGTAT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAAAGATCTGCTGCGCAACCGCGCAAAACCTGGGAA 1920
DB 1861 GAAGCGTTATCGCTGATTTATAAAGATCTGCTGCGCAACCGCGCAAAACCTGGGAA 1920

QY 1921 GAGATCCGCGGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
DB 1921 GAGATCCGCGGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTCCTGCTGACGGGGTTATGCGGTTCAAG 2040
DB 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTCCTGCTGACGGGGTTATGCGGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTTAAAGACGTTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
DB 2041 TATGAAAAACGCAAGTACGACATTTAAAGACGTTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGCTGTTGACCTGATTTAAAAACAAACACATGATGACACCGGATTC 2160
DB 2101 GGTCTGACCTTCTGCTGTTGACCTGATTTAAAAACAAACACATGATGACACCGGATTC 2160
QY 2161 TCCATCGCAAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCGCGCTGG 2220
DB 2161 TCCATCGCAAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCGCGCTGG 2220
QY 2221 GCATGCTCAACATCGACACAGCAAAAGTGAATTTATGGTGAACGGTACTGCGGACCTTC 2280
DB 2221 GCATGCTCAACATCGACACAGCAAAAGTGAATTTATGGTGAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGCTCAACATCGACACAGCAAAAGTGAATTTATGGTGAACGGTACTGCGGACCTTC 2340
DB 2281 AAGGCTCAACATCGACACAGCAAAAGTGAATTTATGGTGAACGGTACTGCGGACCTTC 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGATGAAGGTTCTG 2400
DB 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGATGAAGGTTCTG 2400
QY 2401 GAAGCGGTTTAAAGAACCAACCGCTGGGTGCGTAGCGCTGAAGTTTACGAGGAAGAG 2460
DB 2401 GAAGCGGTTTAAAGAACCAACCGCTGGGTGCGTAGCGCTGAAGTTTACGAGGAAGAG 2460
QY 2461 TTGGCAAAAGATCCAGTATTTGCGCCACCATCGAAAAACCGCCAGAAAGGTGAAATCATG 2520
DB 2461 TTGGCAAAAGATCCAGTATTTGCGCCACCATCGAAAAACCGCCAGAAAGGTGAAATCATG 2520
QY 2521 CCGAAACATCCGAGATGTCGCTTTTCTGGTATGCGCTGCTACTGCGGTGATCAAGGCC 2580
DB 2521 CCGAAACATCCGAGATGTCGCTTTTCTGGTATGCGCTGCTACTGCGGTGATCAAGGCC 2580
QY 2581 GCAGCGGCTGCTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
DB 2581 GCAGCGGCTGCTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAAATTAACAAATTAACAAACCTCGGATCGAGGGAAGGATTTTCAGAAATTC 2700
DB 2641 AACAAACAAACAAATTAACAAATTAACAAACCTCGGATCGAGGGAAGGATTTTCAGAAATTC 2700

RESULT 3

US-09-265-617B-1
; Sequence 1, Application US/09265617B
; Patent No. 6372883
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R.
; APPLICANT: Hurst, David N.
; APPLICANT: Jones, Philip S.
; APPLICANT: Kay, Paul B.
; APPLICANT: Raynham, Tony M.
; APPLICANT: Wilson, Francis X.
; TITLE OF INVENTION: Antiviral Medicaments
; FILE REFERENCE: 20052 antiviral medicaments
; CURRENT APPLICATION NUMBER: US/09/265,617B
; CURRENT FILING DATE: 1999-03-10
; PRIOR APPLICATION NUMBER: GB9806815.8
; PRIOR FILING DATE: 1998-03-30
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1

; LENGTH: 7475
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; INFORMATION: Description of Artificial Sequence: Vector and
; OTHER INFORMATION: Gene Fragments
US-09-265-617B-1

Query Match 81.8%; Score 2700; DB 3; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CCGACACATCGAATGGTGCAGAAACCTTTCGGCGGTATGGCATGATAGGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAGAAACCTTTCGGCGGTATGGCATGATAGGCCCGGAAGAGA 60
QY 61 GTCAATTACAGGTGTGAATGTGAACACAGTAACAGTTATACGATGTCGCAGAGTATGCCG 120
Db 61 GTCAATTACAGGTGTGAATGTGAACACAGTAACAGTTATACGATGTCGCAGAGTATGCCG 120
QY 121 GTGTCCTTTATCAGACCGGTTTCCCGCGTGGTGAACACGCGCACGCGTTTCTGCGAAAA 180
Db 121 GTGTCCTTTATCAGACCGGTTTCCCGCGTGGTGAACACGCGCACGCGTTTCTGCGAAAA 180
QY 181 CGCGGAAAAAGTGAAGCGGCGATGGCGAGCTGAAATTACATTCGCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGGCGATGGCGAGCTGAAATTACATTCGCCAACCGCGTGGCAC 240
QY 241 AACAACTCGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300
Db 241 AACAACTCGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300
QY 301 AGCGCCGTCGCAAAATGTTCGGCGGATTAATAATCTCGCGCGGATCAATTCGGGTGCCAGCG 360
Db 301 AGCGCCGTCGCAAAATGTTCGGCGGATTAATAATCTCGCGCGGATCAATTCGGGTGCCAGCG 360
QY 361 TGGTGGTTCGATGTAGAACGAAGCGGCGTGAAGCCGTGTGAAGCGGCGGTGCACAATC 420
Db 361 TGGTGGTTCGATGTAGAACGAAGCGGCGTGAAGCCGTGTGAAGCGGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATCTATCCGTCGGATGACCAAGGATGCCA 480
Db 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATCTATCCGTCGGATGACCAAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTGATGTCCTGACCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTGATGTCCTGACCAGA 540
QY 541 CACCCATCAACAGTATTTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
QY 601 TGGTGCATTTGGTCAACGCAAAATCGCGCTGTAGCGGCGCCATTAAGTTCTGCTCGG 560
Db 601 TGGTGCATTTGGTCAACGCAAAATCGCGCTGTAGCGGCGCCATTAAGTTCTGCTCGG 560
QY 661 CGCGTCTGGCTCTGGCTGGCTGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGGCTCTGGCTGGCTGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAAACGGGAAGCGACTGGAGTGCATGTCGCGTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAAACGGGAAGCGACTGGAGTGCATGTCGCGTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCATCTGGATGCTGTTGCCAAGATCAGATGGCGCTGGGGCGAA 840
Db 781 ATGAGGGCATCGTTCCCATCTGGATGCTGTTGCCAAGATCAGATGGCGCTGGGGCGAA 840
QY 841 TCGCGCCATTACCGAGTCCGGGCTCGCGGTTGGTGGCGATATCTCGGTAGTGGGATCG 900
Db 841 TCGCGCCATTACCGAGTCCGGGCTCGCGGTTGGTGGCGATATCTCGGTAGTGGGATCG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCCATCAAAAGGATTTTC 960

Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCCATCAAAAGGATTTTC 960
QY 961 GCCTGCTGGGGCAAAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
Db 961 GCCTGCTGGGGCAAAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCGCTCTCATCTGTTGAAAGAAACCAACCTTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCCGCTCTCATCTGTTGAAAGAAACCAACCTTGGCGGCCAATA 1080
QY 1081 CGCAACCGCTCTCCCGCGGTTGGCGAATCAATTAATGAGTGGCAACAGAGGTTT 1140
Db 1081 CGCAACCGCTCTCTCCCGCGGTTGGCGAATCAATTAATGAGTGGCAACAGAGGTTT 1140
QY 1141 CCCGACTTGGAAAGCGGCGAGTGGAGCGCAACCGCAATTAATGAGTGGCAACAGAGGTT 1200
Db 1141 CCCGACTTGGAAAGCGGCGAGTGGAGCGCAACCGCAATTAATGAGTGGCAACAGAGGTT 1200
QY 1201 GCACAAATCTCATGTTTGAACAGCTTATCATCGACTGCAAGCGTGCACCAATGCTTCGCGG 1260
Db 1201 GCACAAATCTCATGTTTGAACAGCTTATCATCGACTGCAAGCGTGCACCAATGCTTCGCGG 1260
QY 1261 TCAGGACGCAATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAATCACTGCATTAATTCG 1320
Db 1261 TCAGGACGCAATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAATCACTGCATTAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACTCCCGTCTTGATTAATGTTTTCGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTCTTGATTAATGTTTTCGCGCGACATCAATACGGTT 1380
QY 1381 CTGGCAATATCTGAAATGAGCTGTTGAOAAATTAATCATCGGCTCGTAAATGCTGTGGA 1440
Db 1381 CTGGCAATATCTGAAATGAGCTGTTGAOAAATTAATCATCGGCTCGTAAATGCTGTGGA 1440
QY 1441 ATTGTAGCGGTAACAAATTTTCAACAGGAACAGCAGTCCGTTTAGGTGTTTTCACGA 1500
Db 1441 ATTGTAGCGGTAACAAATTTTCAACAGGAACAGCAGTCCGTTTAGGTGTTTTCACGA 1500
QY 1501 GCATTTCAACAAAGGACCATAGATTATGAAACTGAAAGAGGTAAATCTGGTAATCTGG 1560
Db 1501 GCATTTCAACAAAGGACCATAGATTATGAAACTGAAAGAGGTAAATCTGGTAATCTGG 1560
QY 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAGAAAT 1620
Db 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAGAAAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGCGCTGACATTAATCTTGGGCAACAGCAGCGCTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCGCTGACATTAATCTTGGGCAACAGCAGCGCTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCGGACAAAGCGTTCAGAGCAAGCTGAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCGGACAAAGCGTTCAGAGCAAGCTGAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTAGCTTTACAAACCGCAAGCTGATCCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTAGCTTTACAAACCGCAAGCTGATCCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGTGTGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGTGTGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGGCTGATGCTGCTGAGCGGGGTTATCGGTTCAAG 2040

|||||
784 GAGGCATCGTTCACCATGCGATGCTGGTGGCAACGATCAGATGCGCTGGCGCAATG 725
QY CGCGCATTTACCGAGTCGCGGCTGCGCGTTGGTGGGATATCTCGGTAGTGGGATACGAC 902
Db CGCGCATTTACCGAGTCGCGGCTGCGCGTTGGTGGGATATCTCGGTAGTGGGATACGAC 665
QY GATACCGAAGACAGCTCATGTTATATCCGCGGTTAAACACCATCAAAACAGATTTTCG 962
Db GATACCGAAGACAGCTCATGTTATATCCGCGGTTCAACCAACCATCAAAACAGATTTTCG 605
QY CTGCTGGGCAAAACAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGGCCAGCGGTGAAG 1022
Db CTGCTGGGCAAAACAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGGCCAGCGGTGAAG 545
QY GGCAATCAGCTGTTGCCCGTCTCACTGGTGGAAAGAAAACACCCCTGGCGCCCAATACG 1082
Db GGCAATCAGCTGTTGCCCGTCTCACTGGTGGAAAGAAAACACCCCTGGCGCCCAATACG 485
QY CAAACCGCTCTCCCGCGCGTGGCCGATTCATTAATGCGAGCTGGCAGCAGGTTTCC 1142
Db CAAACCGCTCTCCCGCGCGTGGCCGATTCATTAATGCGAGCTGGCAGCAGGTTTCC 425
QY CGACTGGAAAGCGGCAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTCATTAGGC 1202
Db CGACTGGAAAGCGGCAGTGAGCGCAACGCAATTAATGAGTTAGCTCACTCATTAGGC 376
QY ACAATTTCTCATGTTTGACAGCTTATCATCGACTGACCGCGTGACCAATGCTTCTGGCGTC 1262
Db GAATTTGATCTGGTTTGACAGCTTATCATCGACTGACCGTGACCAATGCTTCTGGCGTC 316
QY AGGAGCCATCGGAAGCTGTGGTATGGCTGTGCGAGTGTAAATCACTGCATAAATTCGTG 1322
Db AGGAGCCATCGGAAGCTGTGGTATGGCTGTGCGAGTGTAAATCACTGCATAAATTCGTG 256
QY TCGCTCAAGGGCGCACTCCGCTTCTCGATAAGTGTGCGCGGACATCAACCGGTTCT 1382
Db TCGCTCAAGGGCGCACTCCGCTTCTCGATAAGTGTGCGCGGACATCAACCGGTTCT 196
QY GGCAATATTTCTGAATGAGCTGTTGACAAATTAATCAT- CGGCTCGTATAATGTTGGAA 1441
Db GGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCAATCATCGGCTCGTATATGTTGGAA 136
QY TTGTGAGCGGATAACAATTTTCAACAGGAAACAGCC 1477
Db TTGTGAGCGGATAACAATTTTCAACAGGAAACAGAC 100

RESULT 5

US-09-640-882-2
; Sequence 2, Application US/09640882
; Patent No. 6720142
; GENERAL INFORMATION:
; APPLICANT: Hall, Barry G.
; TITLE OF INVENTION: METHOD OF DETERMINING EVOLUTIONARY POTENTIAL OF MUTANT
; TITLE OF INVENTION: RESISTANCE GENES AND USE THEREOF TO SCREEN FOR DRUG
; TITLE OF INVENTION: EFFICACY
; FILE REFERENCE: 176/60851
; CURRENT APPLICATION NUMBER: US/09/640,882
; CURRENT FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/149,813
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 5201
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid
; OTHER INFORMATION: PACSEQ
US-09-640-882-2

Query Match 41.3%; Score 1362; DB 4; Length 5201;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1380; Conservative 0; Mismatches 10; Indels 2; Gaps 1;
QY 3 GACACCATCGAATCGTGCAAAACCTTTTCGCGGTATGGCATATAGCGCCGCGGAAGAGT 62
Db GACACCATCGAATCGTGCGGCAAAACCTTTTCGCGGTATGGCATATAGCGCCGCGGAAGAGT 925
QY 63 CAATTCAGGGTGGTGAATGTGAACACCAAGTAAACGTTATACGATGTGCGAGAGTATGCCGT 122
Db CAATTCAGGGTGGTGAATGTGAACACCAAGTAAACGTTATACGATGTGCGAGAGTATGCCGT 985
QY 123 GTCTCTTATCAGACCGTTTCCGCGGTGGTGAACAGCGCCAGCCAGCTTCTGCGAAACG 182
Db GTCTCTTATCAGACCGTTTCCGCGGTGGTGAACAGCGCCAGCCAGCTTCTGCGAAACG 1045
QY 183 CGGGAAGGTTGGAAGCGGCGATGGCGAGCTGAATTAATATTCACCAACCGGTGGCAAA 242
Db CGGGAAGGTTGGAAGCGGCGATGGCGAGCTGAATTAATATTCACCAACCGGTGGCAAA 1105
QY 243 CAATCGCGGCAACAGTCTGATTGGCGTGGCCACTCCAGTCTGGCCCTGCAC 302
Db CAATCGCGGCAACAGTCTGATTGGCGTGGCCACTCCAGTCTGGCCCTGCAC 1165
QY 303 CGCGCTGCGAAATTTGTCGCGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCGTG 362
Db CGCGCTGCGAAATTTGTCGCGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCGTG 1225
QY 363 GTGGTGTGATGAGAACGCGGCGTGAAGCGCTGTAAGCGGCGGTGCACAACTTT 422
Db GTGGTGTGATGAGAACGCGGCGTGAAGCGCTGTAAGCGGCGGTGCACAACTTT 1285
QY 423 CTCGCGCAACCGTCAGTGGGCTGATCAATTAATCACTATCCGCTGGATGACAGGATGCCATT 482
Db CTCGCGCAACCGTCAGTGGGCTGATCAATTAATCACTATCCGCTGGATGACAGGATGCCATT 1345
QY 483 CTTGTGGAGTCTGCTGCACTAAATGTTCCGCGGTATTTCTTGATGTCTCTGACAGACA 542
Db CTTGTGGAGTCTGCTGCACTAAATGTTCCGCGGTATTTCTTGATGTCTCTGACAGACA 1405
QY 543 CCCATCAACAGTATTATTTCTCCATGAAGACGGTACGCGCTGGGGGTGGAGCATCTG 602
Db CCCATCAACAGTATTATTTCTCCATGAAGACGGTACGCGCTGGGGGTGGAGCATCTG 1465
QY 603 GTCGATTGGGTCAACAGCAAAATCGCGTGTAGCGGCGCCATTAAGTTCTGTCTCGGCG 662
Db GTCGATTGGGTCAACAGCAAAATCGCGTGTAGCGGCGCCATTAAGTTCTGTCTCGGCG 1525
QY 663 CGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
Db CGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 1585
QY 723 GAAAGGGAAGCGACTGGAGTGCATGTCGCGTTTCAACAAACCATGCAAAATGCTGAAT 782
Db GAAAGGGAAGCGACTGGAGTGCATGTCGCGTTTCAACAAACCATGCAAAATGCTGAAT 1645
QY 783 GAGGTCATCGTTCCTCCATCTGCGATGCTGGTTGCGCAACGATCAGATGGCGCTGGGCGCAATG 842
Db GAGGTCATCGTTCCTCCATCTGCGATGCTGGTTGCGCAACGATCAGATGGCGCTGGGCGCAATG 1705
QY 843 CGCGCATTAACCGAGTCTGGCGCTGGCGGATATCTCGGTAGTGGGATACGAC 902
Db CGCGCATTAACCGAGTCTGGCGCTGGCGGATATCTCGGTAGTGGGATACGAC 1765
QY 903 GATACCGAAGACAGCTCATGTTATATCCGCGGTAAACCAACCATCAAAACAGATTTTCG 962
Db GATACCGAAGACAGCTCATGTTATATCCGCGGTAAACCAACCATCAAAACAGATTTTCG 1825
QY 963 CTGCTGGGCGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCGCAGCGGTGAAG 1022
Db CTGCTGGGCGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCGCAGCGGTGAAG 1885
QY 1023 GGCAATCAGCTGTTGCCCGTCTCTACTGTGAAAGAAAAACCAACCTGGCGCCCAATACG 1082

Db 1886 GGCATCAGCTGTGCGCGTCTCACTGGTGAAGAAAAACACCTGCGCCCAATAGC 1945
Qy 1083 CAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGACGTGGCAGCAGAGTTTCC 1142
Db 1946 CAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGACGTGGCAGCAGAGTTTCC 2005
Qy 1143 CGACTGGAAGGGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTAGGC 1202
Db 2006 CGACTGGAAGGGGCGAGTGAGCGCAACGCAATTAATGTAGTTAG--CGCAATTGATC 2063
Qy 1203 ACAATCTCATGTTGACAGCTTATCATCGACTGCAACGCGTGACCAATGCTTCTGCGCTC 1262
Db 2064 TGAATTTCTCATGTTTACAGCTTATCATCGACTGCAACGCGTGACCAATGCTTCTGCGCTC 2123
Qy 1263 AGCAGCCATCGAAGCTGTGTGTATGGCTGTGCGAGTGTGTAATCACTGCAATTAATTCGTG 1322
Db 2124 AGCAGCCATCGAAGCTGTGTGTATGGCTGTGCGAGTGTGTAATCACTGCAATTAATTCGTG 2183
Qy 1323 TCGCTCAAGGCGCACTCCGTTCTCGATAATGTTTTTGGCGCGACATCAATACGTTTCT 1382
Db 2184 TCGCTCAAGGCGCACTCCGTTCTCGATAATGTTTTTGGCGCGACATCAATACGTTTCT 2243
Qy 1383 GGCATATTTCT 1394
Db 2244 GGCATATTTCT 2255

RESULT 6

US-09-640-882-3
; Sequence 3, Application US/09640882
; Patent No. 6720142
; GENERAL INFORMATION:
; APPLICANT: Hall, Barry G.
; TITLE OF INVENTION: METHOD OF DETERMINING EVOLUTIONARY POTENTIAL OF MUTANT
; TITLE OF INVENTION: RESISTANCE GENES AND USE THEREOF TO SCREEN FOR DRUG
; TITLE OF INVENTION: EFFICACY
; FILE REFERENCE: 176/60851
; CURRENT APPLICATION NUMBER: US/09/640,882
; CURRENT FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/149,813
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 5201
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid
US-09-640-882-3

Query Match 41.3%; Score 1362; DB 4; Length 5201;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1380; Conservative 0; Mismatches 10; Indels 2; Gaps 1;

Qy 3 GACACCATCGAATGGTGCAGAAACCTTTTCGCGGTATGGCATATAGCGCCGGAAGAGAGT 62
Db 866 GACACCATCGAATGGCGCAAAACCTTTTCGCGGTATGGCATATAGCGCCGGAAGAGAGT 925
Qy 63 CAATTCAGGGTGGTAATGTGAACACCACTAGCTTATACGATTCGCGAGATATGCCGGT 122
Db 926 CAATTCAGGGTGGTAATGTGAACACCACTAGCTTATACGATTCGCGAGATATGCCGGT 985
Qy 123 GTCTCTTATCAGACGTTTCCCGGTGGTGAACCAAGCCAGCCACGTTTCTGCGAAAAACG 182
Db 986 GTCTCTTATCAGACGTTTCCCGGTGGTGAACCAAGCCAGCCACGTTTCTGCGAAAAACG 1045
Qy 183 CGGAAAAAGTGAAGCGGCGATCGCGAGCTGTAATTAATTCACAAACCGCGTGGCAAA 242
Db 1046 CGGAAAAAGTGAAGCGGCGATCGCGAGCTGAATTAATTCACAAACCGCGTGGCAAA 1105

Qy 243 CAACCTGGCGGCAAAACAGTCTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTTGAC 302
Db 1106 CAACCTGGCGGCAAAACAGTCTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTTGAC 1165
Qy 303 GCGCGCTCGCAAAATTTGTCGCGCGGATTAATCTCGCGCGGATCAATCGGTTGCCAGGCTG 362
Db 1166 GCGCGCTCGCAAAATTTGTCGCGCGGATTAATCTCGCGCGGATCAATCGGTTGCCAGGCTG 1225
Qy 363 GTGCTGTCATGTTAGTAAAGCGGCTGAAAGCGGCTGTAAGCGGCTGCAAAATCTT 422
Db 1226 GTGCTGTCATGTTAGTAAAGCGGCTGAAAGCGGCTGTAAGCGGCTGCAAAATCTT 1285
Qy 423 CTCGCGCAACGCGTCAGTGGCTGATCAATTAATCTCGCTGGATGACAGGATGCCATT 482
Db 1286 CTCGCGCAACGCGTCAGTGGCTGATCAATTAATCTCGCTGGATGACAGGATGCCATT 1345
Qy 483 GCTGTGAAAGCTGCTGCACCTAAATGTTCCGCGGTTATTTCTTGATGTTCTCTGACGACGA 542
Db 1346 GCTGTGAAAGCTGCTGCACCTAAATGTTCCGCGGTTATTTCTTGATGTTCTCTGACGACGA 1405
Qy 543 CCNATCAACAGTATTTTCTCCATGAAGACGTTACGCGATGCGCGTGGAGCATCTG 602
Db 1406 CCCATCAACAGTATTTTCTCCATGAAGACGTTACGCGATGCGCGTGGAGCATCTG 1465
Qy 603 GTCGATTGGTCAACAGCAATCGCGCTGTAGCGGCGCCATTAAAGTTCTGTCTCGCG 662
Db 1466 GTCGATTGGTCAACAGCAATCGCGCTGTAGCGGCGCCATTAAAGTTCTGTCTCGCG 1525
Qy 663 GGTCTGCTGTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAGCG 722
Db 1526 GGTCTGCTGTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAGCG 1585
Qy 723 GAACGGGAAGCGCACTGGAGTGCATGTCGCGTTTCAACAAACCATGCAAAATGCTGAAT 782
Db 1586 GAACGGGAAGCGCACTGGAGTGCATGTCGCGTTTCAACAAACCATGCAAAATGCTGAAT 1645
Qy 783 GAGGCGATCTGCTCCACTGCGATGCTGTTGCAACAGCATCAGATGCGGCTGGCGCAATG 842
Db 1646 GAGGCGATCTGCTCCACTGCGATGCTGTTGCAACAGCATCAGATGCGGCTGGCGCAATG 1705
Qy 843 GCGCGCATTTACCGAGTCCGCGCTGCGGTTGGTGGGATATCTCGGTAGTGGATACGAC 902
Db 1706 GCGCGCATTTACCGAGTCCGCGCTGCGGTTGGTGGGATATCTCGGTAGTGGATACGAC 1765
Qy 903 GATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTCG 962
Db 1766 GATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTCG 1825
Qy 963 CTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGAAG 1022
Db 1826 CTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGAAG 1885
Qy 1023 GGCATACAGCTGTGCGCGCTCTCACTGGTGAAGAAAAACCAACCTTGGCGCCCAATAGC 1082
Db 1886 GGCATACAGCTGTGCGCGCTCTCACTGGTGAAGAAAAACCAACCTTGGCGCCCAATAGC 1945
Qy 1083 CAAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAATGACGTGCAACAGAGTTTTC 1142
Db 1946 CAAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAATGACGTGCAACAGAGTTTTC 2005
Qy 1143 CGACTGGAAGGGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTAGGC 1202
Db 2006 CGACTGGAAGGGGCGAGTGAGCGCAACGCAATTAATGTAGTTAG--CGCAATTGATC 2063
Qy 1203 ACAATTTCTCATGTTTACAGCTTATCATCGACTGCAACGCGTGACCAATGCTTCTGCGCTC 1262
Db 2064 TGAATTTCTCATGTTTACAGCTTATCATCGACTGCAACGCGTGACCAATGCTTCTGCGCTC 2123
Qy 1263 AGCAGCCATCGAAGCTGTGTGTATGGCTGTGCGAGTGTGTAATCACTGCAATTAATTCGTG 1322
Db 2124 AGCAGCCATCGAAGCTGTGTGTATGGCTGTGCGAGTGTGTAATCACTGCAATTAATTCGTG 2183
Qy 1323 TCGCTCAAGGCGCACTCCGTTCTCGATAATGTTTTTGGCGCGACATCAATACGTTTCT 1382

|||||
2184 TCGCTCAAGGCGCACTCCGTTCTGGATAATGTTTTTGGCCCGACATCAATACGGTTCT 2243
Qy
1383 GCGAAATATTTCT 1394
Db
2244 GCGAAATATTTCT 2255

RESULT 7
US-08-778-717-5
; Sequence 5, Application US/08778717
; Patent No. 6602689
; GENERAL INFORMATION:
; APPLICANT: UENO, EIICHI
; APPLICANT: NOBUYUKI, FUJII
; APPLICANT: OKADA, NASHISA
; TITLE OF INVENTION: FUSED DNA SEQUENCE, FUSED PROTEIN
; TITLE OF INVENTION: EXPRESSED FROM SAID FUSED DNA SEQUENCE AND METHOD FOR
; TITLE OF INVENTION: EXPRESSING SAID DNA SEQUENCE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/778,717
; FILING DATE: 12-DEC-1996
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 352225/1995
; FILING DATE: 28-DEC-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2084-031-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-413-3000
; TELEFAX: 703-413-2220
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4557 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; ORIGINAL SOURCE:
; ORGANISM: E. COLI
; STRAIN: BL21 (DE3)
; PUBLICATION INFORMATION:
; AUTHORS: NOBUYUKI FUJII ET AL,
; TITLE: FUSED DNA SEQUENCE, FUSED PROTEIN EXPRESSED
; TITLE: FROM SAID FUSED DNA SEQUENCE AND METHOD OF
; TITLE: EXPRESSING SAID FUSED PROTEIN
; RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 TO 4557
US-08-778-717-5

Query Match 36.5%; Score 1204; DB 4; Length 4557;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCGACACCATCAATGGTGCAGAAACCTTTCCGGGTATGCGATGATAGCGCCGGAAGAGA 60
Db 2563 CCGACACCATCAATGGTGCAGAAACCTTTCCGGGTATGCGATGATAGCGCCGGAAGAGA 2622

61 GTCAATTACGGTGGTGAATGCTGAAACCCAGTAACGTTATACGATGTCGACAGTATGCCG 120
2623 GTCAATTACGGTGGTGAATGCTGAAACCCAGTAACGTTATACGATGTCGACAGTATGCCG 2682
Qy 121 GTGTCTCTTATCAGACCGGTTTCCCGCGTGGTGAACACGAGCCAGCCACCGTTTCTCGGAAA 180
Db 2683 GTGTCTCTTATCAGACCGGTTTCCCGCGTGGTGAACACGAGCCAGCCACCGTTTCTCGGAAA 2742
Qy 181 CGCGGAAAAAGTGGAAAGCGCGGATGCGGAGCTGAAATTAATTCCTCCCAACCGCTGGCAC 240
2743 CGCGGAAAAAGTGGAAAGCGCGGATGCGGAGCTGAAATTAATTCCTCCCAACCGCTGGCAC 2802
Qy 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATGCGGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 2803 AACAACTGGCGGGCAAAACAGTCGTTGCTGATGCGGTTGCCACCTCCAGTCTGGCCCTGC 2862
Qy 301 ACGCGCGCTCGCAAAATGTCGCGCGGATTAATCTCGCGCGGATCACTCGGCTGGTGCACG 360
Db 2863 ACGCGCGCTCGCAAAATGTCGCGCGGATTAATCTCGCGCGGATCACTCGGCTGGTGCACG 2922
Qy 361 TGGTGGTGTGATGGTAGAACGAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCAACAATC 420
2923 TGGTGGTGTGATGGTAGAACGAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCAACAATC 2982
Qy 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCACTAACTATCCGCTGGATACCCAGGATGCCA 480
2983 TTCTCGCGCAACGCGTCAGTGGGCTGATCACTAACTATCCGCTGGATACCCAGGATGCCA 3042
Qy 481 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 540
3043 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA 3102
Qy 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGAGCGGTACGCGACTGGGGCTGGAGCATC 600
3103 CACCCATCAACAGTATTAATTTCTCCCATGAAGAGCGGTACGCGACTGGGGCTGGAGCATC 3162
Qy 601 TGGTGGCATTTGGGTACCAGCAAAATCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCTCG 660
3163 TGGTGGCATTTGGGTACCAGCAAAATCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCTCG 3222
Qy 661 CGGCTCTCGCTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
3223 CGGCTCTCGCTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 3282
Qy 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
3283 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 3342
Qy 781 ATGAGGGCATCGTTCCCACTCGGATGCTGGTTGCCAACGATCAGATGGCGCTGGCGGCA 840
3343 ATGAGGGCATCGTTCCCACTCGGATGCTGGTTGCCAACGATCAGATGGCGCTGGCGGCA 3402
Qy 841 TGGCGCCATTACCGAGTCCGGGCTGGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
3403 TGGCGCCATTACCGAGTCCGGGCTGGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 3462
Qy 901 ACGATACCGGAAGACAGCTCATGTTATATCCGCGCTTAACCAACCATCAAAAGGATTTTC 960
3463 ACGATACCGGAAGACAGCTCATGTTATATCCGCGCTTAACCAACCATCAAAAGGATTTTC 3522
Qy 961 GCCTCTGGGGCAAAACAGCGTGGAGCCGCTTGTGCGAACTCTCTCAGGCGCAGGCGGTGA 1020
3523 GCCTCTGGGGCAAAACAGCGTGGAGCCGCTTGTGCGAACTCTCTCAGGCGCAGGCGGTGA 3582
Qy 1021 AGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAACCAACCTGGCGGCCCAATA 1080
3583 AGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAACCAACCTGGCGGCCCAATA 3642
Qy 1081 CGCAAAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAATGCACTGGCAGCAGCAGGTTT 1140
3643 CGCAAAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAATGCACTGGCAGCAGCAGGTTT 3702
Qy 1141 CCCGACTGGAAAGCGGGCAGTGAGGCCAACGCAATTAATGTGAGTTAGTCTACTCATTTAG 1200

```

Db      3703 CCCGACTGGAAGGGCGGAGTGGAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTAG 3762
Qy      1201 GCAC 1204
Db      3763 GCAC 3766

RESULT 8
US-08-148-675A-2
; Sequence 2, Application US/08148675A
; Patent No. 5506121
; GENERAL INFORMATION:
; APPLICANT: Skerra, Arne; Schmidt, Thomas
; TITLE OF INVENTION: FUSION PEPTIDES WITH BINDING ACTIVITY FOR
; TITLE OF INVENTION: STREPTAVIDIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/148,675A
; FILING DATE: 3-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 42 37 113.9
; FILING DATE: 03-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Tsai, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: HUBR 1041
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3832 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-148-675A-2

Query Match          36.4%; Score 1200.8; DB 1; Length 3832;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1202; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy      1 CGGACACCATCGAATCGGTGCAAAACCTTTTCGGCGTATGCATGATAGCGCCGCGAAGAGA 60
Db      3 CGGACACCATCGAATGGCGCAAAACCTTTTCGGCGTATGCATGATAGCGCCGCGAAGAGA 62

Qy      61 GTCAATTACAGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGACAGTATGCCG 120
Db      63 GTCAATTACAGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGACAGTATGCCG 122

Qy      121 GTGTCTCTTTATCAGACCGGTTTCCCGCGTGGTGAAACGAGGCCAGCCACGTTTCTGCGAAAA 180
Db      123 GTGTCTCTTTATCAGACCGGTTTCCCGCGTGGTGAAACGAGGCCAGCCACGTTTCTGCGAAAA 182

Qy      181 CCGCGGAAAAAGTGGAAAGCGGCGATGCGGAGACTGAATTACATTCCTCCCAACCGCGTGGCAC 240
Db      183 CCGCGGAAAAAGTGGAAAGCGGCGATGCGGAGACTGAATTACATTCCTCCCAACCGCGTGGCAC 242

Qy      241 AACAACTGCGCGGCAAAACAGTTCGTTGCTGATTGGGTTGCCACCTCCAGTCTGGCCCTTGC 300

```

APPLICANT: Matis, Louis A.
APPLICANT: Mueller, Bileen Elliott
APPLICANT: Nye, Steven H.
APPLICANT: Rollins, Scott
APPLICANT: Rother, Russell P.
APPLICANT: Springhorn, Jeremy P.
APPLICANT: Squinto, Stephen C.
APPLICANT: Thomas, James C.
APPLICANT: Wilkins, James A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT
OF INFLAMMATORY DISEASES
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seth A. Fidel
CITY: New Haven
STATE: Connecticut
COUNTRY: USA
ZIP: 06511
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.4Mb storage
COMPUTER: Macintosh Cetris 610
OPERATING SYSTEM: System 7
SOFTWARE: WordPerfect 3.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,283A
FILING DATE: June 7, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/236,208
FILING DATE: 02-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Seth A. Fidel.
REGISTRATION NUMBER: 38,449
REFERENCE/DOCKET NUMBER: ALX-152.1 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203)776-1790
TELEFAX: (203)772-3655
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 5248 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Circular
MOLECULE TYPE: Other nucleic acid
DESCRIPTION: PET Trc SOS/NI
DESCRIPTION: prokaryotic expression vector
US-08-487-283A-18

Query Match 36.3%; Score 1199.2; DB 3; Length 5248;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGCRAAACTTTTCGGGTATGGCATAGAGCCCGGAAGAGA 60
DB 4810 CGGACACCATCGAATGGCGCAAAACCTTTCGGGTATGGCATAGAGCCCGGAAGAGA 4751
QY 61 GTCMAATTCAGGGTGGTGAATGTGAACACAGTAACCGTTATACGATCTCGCAGAGTATGCCG 120
DB 4750 GTCMAATTCAGGGTGGTGAATGTGAACACAGTAACCGTTATACGATCTCGCAGAGTATGCCG 4691
QY 121 GTGTCTTATCAGACCGTTTCCTCCCGGTGGTGAACACAGCCAGCCACGTTTCGCGAAAA 180
DB 4690 GTGTCTTATCAGACCGTTTCCTCCCGGTGGTGAACACAGCCAGCCACGTTTCGCGAAAA 4631
QY 181 CGCGGAAAGTGAAGCGGGATGGCGGAGCTGAATTACATTCCTCCACCGCGTGGCAC 240
DB 4630 CGCGGAAAGTGAAGCGGGATGGCGGAGCTGAATTACATTCCTCCACCGCGTGGCAC 4571
QY 241 AACAACTCGCGGGCAAAACAGTCTGTTCGATTTGGGGTTGCCACCTCCAGTCTGGCCCTGC 300
DB 4570 AACAACTCGCGGGCAAAACAGTCTGTTCGATTTGGGGTTGCCACCTCCAGTCTGGCCCTGC 4511
QY 301 ACGCGCCGTGCGCAAAATGTTCGCGCGATTAATAATCTCGCGCGATCAACTGGGTGCCAGCG 360

DB 4510 ACGCGCCGTGCGCAAAATGTTCGCGCGATTAATAATCTCGCGCGATCAACTGGGTGCCAGCG 4451
QY 361 TGGTGGTGTTCGATGTAGAACGAAGCGCGCTGGAAGCCTGTAAAGCGCGCTGCACAATC 420
DB 4450 TGGTGGTGTTCGATGTAGAACGAAGCGCGCTGGAAGCCTGTAAAGCGCGCTGCACAATC 4391
QY 421 TTCTCGGCAACGCGTCACTGGGCTGATCAATACTATCCGCTGGATGACGAGATGCCA 480
DB 4390 TTCTCGGCAACGCGTCACTGGGCTGATCAATACTATCCGCTGGATGACGAGATGCCA 4331
QY 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCGCGGCTTAATTTCTTGATGTCTCGACCA 540
DB 4330 TTGCTGTGGAAGCTGCCTGCACTAATGTTCGCGGCTTAATTTCTTGATGTCTCGACCA 4271
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGCGCTGGAGCATC 600
DB 4270 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGCGCTGGAGCATC 4211
QY 601 TGGTGGCATTTGGGTCAACAGCAAAATCGGCTGTGTAGCGGGCCCAATTAAGTCTGTCTCG 660
DB 4210 TGGTGGCATTTGGGTCAACAGCAAAATCGGCTGTGTAGCGGGCCCAATTAAGTCTGTCTCG 4151
QY 661 CGGCTCTGGCTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
DB 4150 CGGCTCTGGCTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 4091
QY 721 CGGAACGGGAAGCGGACTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAATTCGTGA 780
DB 4090 CGGAACGGGAAGCGGACTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAATTCGTGA 4031
QY 781 ATGAGGGCATCGTTCCCACTCGCATGTGGTTGGCAACGATCAGATGGCGCTGGCGCGAA 840
DB 4030 ATGAGGGCATCGTTCCCACTCGCATGTGGTTGGCAACGATCAGATGGCGCTGGCGCGAA 3971
QY 841 TGGCGGCAATTAACGAGTCCGGCTCGCGTGGTGGCGGATATCTCGGTAGTGGGATAG 900
DB 3970 TGGCGGCAATTAACGAGTCCGGCTCGCGTGGTGGCGGATATCTCGGTAGTGGGATAG 3911
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACCAACCATCAAAAGGATTTTC 960
DB 3910 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACCAACCATCAAAAGGATTTTC 3851
QY 961 GCCTGTGGGGCAAAACAGCGTGCACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
DB 3850 GCCTGTGGGGCAAAACAGCGTGCACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 3791
QY 1021 AGGCAATCAGCTGTGGCGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGGCCAATA 1080
DB 3790 AGGCAATCAGCTGTGGCGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGGCCAATA 3731
QY 1081 CGCAACCGCTCTCCCGCGCTTGGCGGATTCATTAATCAGCTGGCAGCAGAGTTT 1140
DB 3730 CGCAACCGCTCTCCCGCGCTTGGCGGATTCATTAATCAGCTGGCAGCAGAGTTT 3671
QY 1141 CCCGACTGGAAGGGGGCAGTGAGCGCAACCAATTAATGTGAGTTAGTCTCACTCATTTAG 1200
DB 3670 CCCGACTGGAAGGGGGCAGTGAGCGCAACCAATTAATGTGAGTTAGTCTCACTCATTTAG 3611
QY 1201 GCAC 1204
DB 3610 GCAC 3607
RESULT 10
PCT-US96-05611A-21/c
; Sequence 21, Application PC/TUS9605611A
; GENERAL INFORMATION:
; APPLICANT: Mueller, John P.
; APPLICANT: Lenardo, Michael J.
; APPLICANT: McFarland, Henry F.
; APPLICANT: Matis, Louis A.
; APPLICANT: Mueller, Bileen Elliott

APPLICANT: Nye, Steven H.
APPLICANT: Pelfrey, Clara M.
APPLICANT: Squinto, Stephen P.
APPLICANT: Wilkins, James A.
TITLE OF INVENTION: Modified Myelin Protein Molecules
NUMBER OF SEQUENCES: 29
CORRESPONDENCE ADDRESS:
ADDRESS: Maurice M. Klee
STREET: 1951 Burr Street
CITY: Fairfield
STATE: Connecticut
COUNTRY: USA
ZIP: 06430
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 0.8 Mb storage
COMPUTER: Macintosh Centris 610
OPERATING SYSTEM: System 7
SOFTWARE: Microsoft Word 6.0.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/05611A
FILING DATE: 02-MAY-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/431,644
FILING DATE: May 2, 1995
APPLICATION NUMBER: 08/431,648
FILING DATE: May 2, 1995
APPLICATION NUMBER: 08/482,114
FILING DATE: June 7, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Klee, Maurice M.
REGISTRATION NUMBER: 30,399
REFERENCE/DOCKET NUMBER: ALX-129
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 255 1400
TELEFAX: (203) 254 1101
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 5248 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Circular
MOLECULE TYPE: Other nucleic acid
DESCRIPTION: DET Trc S05/NI
DESCRIPTION: Prokaryotic expression vector
PCT-US96-05611A-21

Query Match 36.3%; Score 1199.2; DB 5; Length 5248;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CGGACACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA 60
DB 4810 CGGACACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA 4751
QY 61 GTCGAATTCAGGCGTGTGAATCTGAACACAGTAAAGTTATACGATGTCGAGAGTATCCG 120
DB 4750 GTCGAATTCAGGCGTGTGAATCTGAACACAGTAAAGTTATACGATGTCGAGAGTATCCG 4691
QY 121 GTGCTCTTATCAGACCGTTTCCCGCGTGTGAACACAGGCGCAGCAGCTTCTGCGAATA 180
DB 4690 GTGCTCTTATCAGACCGTTTCCCGCGTGTGAACACAGGCGCAGCAGCTTCTGCGAATA 4631
QY 181 CGCGGGAAGAGTGAACCGCGATGGCGGAGCTGAATTCATCCCAACCGCGTGGCAC 240
DB 4630 CGCGGGAAGAGTGAACCGCGATGGCGGAGCTGAATTCATCCCAACCGCGTGGCAC 4571
QY 241 AACAACTGGCGGCAACAGTCGTTGCTGATTTGGCGTTGACCTCCAGTCCGCTGCGCTGC 300
DB 4570 AACAACTGGCGGCAACAGTCGTTGCTGATTTGGCGTTGACCTCCAGTCCGCTGCGCTGC 4511
QY 301 ACGCGCGTGCAGAAATGTGCGCGGATTAATCTCGCGCGATCAACTCGGTGCCAGCG 360

DB 4510 ACGCGCGTGCAGAAATGTGCGCGGATTAATCTCGCGCGATCAACTCGGTGCCAGCG 4451
QY 361 TGGTGGTGTGATGTAAGACGAAGCGCGCTGGAAGCTCTGAAAGCGCGGTGCACAAATC 420
DB 4450 TGGTGGTGTGATGTAAGACGAAGCGCGCTGGAAGCTCTGAAAGCGCGGTGCACAAATC 4391
QY 421 TTTCTGCGCAACCGCTGAGTGGGCTGATCAATTAATCTAATCGCTGGATGCCAGATGCCA 480
DB 4390 TTTCTGCGCAACCGCTGAGTGGGCTGATCAATTAATCTAATCGCTGGATGCCAGATGCCA 4331
QY 481 TTTCTGTTGGAAGCTGCTGCACTAATGTTCCGGGTTATTTCTTGATGTTCTTGACAGA 540
DB 4330 TTTCTGTTGGAAGCTGCTGCACTAATGTTCCGGGTTATTTCTTGATGTTCTTGACAGA 4271
QY 541 CACCATCAACAGTATTTATTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
DB 4270 CACCATCAACAGTATTTATTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 4211
QY 601 TGGTGGCATTTGGGTACCGCAATTCGGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
DB 4210 TGGTGGCATTTGGGTACCGCAATTCGGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 4151
QY 661 CGGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 4150 CGGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 4091
QY 721 CGGAACGGGAAGGAGTGGAGTCCGATGTCGGTTTTCAACAAACCATGCAATGCTGA 780
DB 4090 CGGAACGGGAAGGAGTGGAGTCCGATGTCGGTTTTCAACAAACCATGCAATGCTGA 4031
QY 781 ATGAGGCGCATGTTCCCACTGCGATCGCTGTTGCCAAGATCAGATGGCGCTGGGCGCAA 840
DB 4030 ATGAGGCGCATGTTCCCACTGCGATCGCTGTTGCCAAGATCAGATGGCGCTGGGCGCAA 3971
QY 841 TGGCGCGCATTTACCGAGTCCGGCTCGCGGTTGGTGGGATATCTCGTAGTGGGATACG 900
DB 3970 TGGCGCGCATTTACCGAGTCCGGCTCGCGGTTGGTGGGATATCTCGTAGTGGGATACG 3911
QY 901 ACGATACCGAAGACAGTCAATGTTATATCCGCGGTTAACCCACCATCAACAGGATTTTC 960
DB 3910 ACGATACCGAAGACAGTCAATGTTATATCCGCGGTTAACCCACCATCAACAGGATTTTC 3851
QY 961 GCCTGCTGGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCAGCGGTGA 1020
DB 3850 GCCTGCTGGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCAGCGGTGA 3791
QY 1021 AGGGCAATCAGCTGTGCGCGTCTCACTGGTGAAGAAAAAACCCCTGGCGCCCAATA 1080
DB 3790 AGGGCAATCAGCTGTGCGCGTCTCACTGGTGAAGAAAAAACCCCTGGCGCCCAATA 3731
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTTCAATGATGAGTGGCGACGAGGTTT 1140
DB 3730 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTTCAATGATGAGTGGCGACGAGGTTT 3671
QY 1141 CCGAGTGGAAAGCGGCGAGTGGCGCAACGCAATTAATGAGTGGTCTCACTCATTTAG 1200
DB 3670 CCGAGTGGAAAGCGGCGAGTGGCGCAACGCAATTAATGAGTGGTCTCACTCATTTAG 3611
QY 1201 GCAC 1204
DB 3610 GCAC 3607

RESULT 11
US-10-263-103-35/c
; Sequence 35, Application US/10263103
; Patent No. 6821723
; GENERAL INFORMATION:
; APPLICANT: AVENTIS PASTEUR
; APPLICANT: Chevallier, Michel
; APPLICANT: El Habib, Raphaelle
; APPLICANT: Krell, Tino
; APPLICANT: Sodoyer, Regis

; TITLE OF INVENTION: gp41 antigen
; FILE REFERENCE: 01-1692-A
; CURRENT APPLICATION NUMBER: US/10/263,103
; CURRENT FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 5312
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: plasmid pET-
us-10-263-103-35

Query Match 36.3%; Score 1199.2; DB 4; Length 5312;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	1	CCGACACCATCGAATGGTGC	AAACCTTTCCGGGTATGGCATGATAGCGCCGGAAGAGA	60
DB	4622	CGGACACCATCGAATGGCG	CAAAACCTTTCCGGGTATGGCATGATAGCGCCGGAAGAGA	4563
QY	61	GTCAATTCAGGGTGGTGAAT	GTGAACACAGTAACCGTTATACGATGTCGACAGTATGCCG	120
DB	4562	GTCAATTCAGGGTGGTGAAT	GTGAACACAGTAACCGTTATACGATGTCGACAGTATGCCG	4503
QY	121	GTGTCCTTATCAGACCGTT	TCCCGCGTGTGAACACGACCGACCGTTCCTCGCAAAA	180
DB	4502	GTGTCCTTATCAGACCGTT	TCCCGCGTGTGAACACGACCGACCGTTCCTCGCAAAA	4443
QY	181	CGCGGAAAAGTGAAGCGG	CGATGGCGGAGCTGAATACATTCGCCAACCGCTGGGAC	240
DB	4442	CGCGGAAAAGTGAAGCGG	CGATGGCGGAGCTGAATACATTCGCCAACCGCTGGGAC	4383
QY	241	AACAACTGCGGGCAAA	CAGTCTGCTGATTGGCGTTGGCCACCTCCAGTCTGGCCCTGC	300
DB	4382	AACAACTGCGGGCAAA	CAGTCTGCTGATTGGCGTTGGCCACCTCCAGTCTGGCCCTGC	4323
QY	301	ACGCGCGTTCGCAAA	TTGTCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG	360
DB	4322	ACGCGCGTTCGCAAA	TTGTCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG	4263
QY	361	TGGTGGTTCGATGTAG	ACGAAGCGGCTGGAAGCGCTGTAAGCGGCGGTGCAATC	420
DB	4262	TGGTGGTTCGATGTAG	ACGAAGCGGCTGGAAGCGCTGTAAGCGGCGGTGCAATC	4203
QY	421	TTCTCGCGCAACGCT	CAGTGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCA	480
DB	4202	TTCTCGCGCAACGCT	CAGTGGCTGATCATTAATCTATCCGCTGGATGACCAAGGATGCA	4143
QY	481	TTGCTGTGGAAGCT	GCCTGCCTAAATGTTTCGCGCGTTATTTCTTGATGTTCTTGACCCAGA	540
DB	4142	TTGCTGTGGAAGCT	GCCTGCCTAAATGTTTCGCGCGTTATTTCTTGATGTTCTTGACCCAGA	4083
QY	541	CACCATCAACAGTAT	TTTCTCCATGAAGACGGTACCGCATCGGCGTGGAGCATC	600
DB	4082	CACCATCAACAGTAT	TTTCTCCATGAAGACGGTACCGCATCGGCGTGGAGCATC	4023
QY	601	TGGTGCATTTGGT	CACAGCAAAATCGCTGTTAGCGGGCCCATTAAGTTCCTGCTCGG	660
DB	4022	TGGTGCATTTGGT	CACAGCAAAATCGCTGTTAGCGGGCCCATTAAGTTCCTGCTCGG	3963
QY	661	CGCGTCTGCGT	CTGGCTGGCATATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	3962	CGCGTCTGCGT	CTGGCTGGCATATAATCTCACTCGCAATCAAAATTCAGCCGATAG	3903
QY	721	CGGAACGGGAACGG	CATGAGTGCCATGTCGCGTTTCAACAAACCATGCAAAATGCTGA	780
DB	3902	CGGAACGGGAACGG	CATGAGTGCCATGTCGCGTTTCAACAAACCATGCAAAATGCTGA	3843
QY	781	ATGAGGGCATCGT	TCCCACTGGATGCTGTTGCCAACGATCAGATGGCGTGGCGCA	840
DB	3842	ATGAGGGCATCGT	TCCCACTGGATGCTGTTGCCAACGATCAGATGGCGTGGCGCA	3783

QY	841	TGCGCGCATTTACGAGT	CCGGGCTGGCGTGGTGGGATATCTCGGTAGTGGATACG	900
DB	3782	TGCGCGCATTTACGAGT	CCGGGCTGGCGTGGTGGGATATCTCGGTAGTGGATACG	3723
QY	901	ACGATACCGAAGACAG	CTCATGTTATATCCCGCGTAAACACCATCAACACGAGATTTC	960
DB	3722	ACGATACCGAAGACAG	CTCATGTTATATCCCGCGTAAACACCATCAACACGAGATTTC	3663
QY	961	GCCTGCTGGGCAAA	ACCAGCGCTGGACCGCTTGTGCAACTCTCTCAGGCGCCGCGTGA	1020
DB	3662	GCCTGCTGGGCAAA	ACCAGCGCTGGACCGCTTGTGCAACTCTCTCAGGCGCCGCGTGA	3603
QY	1021	AGGCAATCAGCTGT	TGGCCCGTCTCACTGTTAAAGAAAAACCCCTGGCGCCCAATA	1080
DB	3602	AGGCAATCAGCTGT	TGGCCCGTCTCACTGTTAAAGAAAAACCCCTGGCGCCCAATA	3543
QY	1081	CGCAACCGCTCT	CCCCCGCGTTGGCGGATTCATTAATCAGCTGGCAGCAGGTTT	1140
DB	3542	CGCAACCGCTCT	CCCCCGCGTTGGCGGATTCATTAATCAGCTGGCAGCAGGTTT	3483
QY	1141	CCCAGCTGAAA	AGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200
DB	3482	CCCAGCTGAAA	AGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	3423
QY	1201	GCAC	1204	
DB	3422	GCAC	3419	

RESULT 12

US-08-929-967-1/c
; Sequence 1, Application US/08929967
; Patent No. 5891637
; GENERAL INFORMATION:
; APPLICANT: Ruppert, Siegfried J.W.
; TITLE OF INVENTION: Construction of Full-length cDNA Libraries
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/929,967
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1035R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5443 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-929-967-1

Query Match 36.3%; Score 1199.2; DB 2; Length 5443;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCGACACATCGAATGGTGCAAAACCTTTCCGGTATGGCATATAGCGCCCGGAAGAGA 60
Db 4810 CGGACACCATCGAATGGCGCAAAACCTTTCCGGTATGGCATATAGCGCCCGGAAGAGA 4751
QY 61 GTCATTTAGGGTGGTGAATGTGAACACAGTAAAGTTATACGATGTCGAGATATGCCG 120
Db 4750 GTCATTTAGGGTGGTGAATGTGAACACAGTAAAGTTATACGATGTCGAGATATGCCG 4691
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACAGGCGCAGCTTTCTGCGAAAA 180
Db 4690 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACAGGCGCAGCTTTCTGCGAAAA 4631
QY 181 CGCGGGAAGAGTGAAGCGCGCATGGCGAGCTGAATACATTCCTCAACCGCGTGGCAC 240
Db 4630 CGCGGGAAGAGTGAAGCGCGCATGGCGAGCTGAATACATTCCTCAACCGCGTGGCAC 4571
QY 241 AACAACTGGCGGGCAACAGCTGCTGTGATGTCGCGTGGCGCTCCACCTCCAGCTTGGCCCTGC 300
Db 4570 AACAACTGGCGGGCAACAGCTGCTGTGATGTCGCGTGGCGCTCCACCTCCAGCTTGGCCCTGC 4511
QY 301 ACGGCGCGTGGCAAAATGTCGCGCGGATTAATCTCGCGCGCATCAACTGGGTGCCAGCG 360
Db 4510 ACGGCGCGTGGCAAAATGTCGCGCGGATTAATCTCGCGCGCATCAACTGGGTGCCAGCG 4451
QY 361 TGGTGTGTGTCGATGAGACGAGCGGCTGAAGCCTGTAAGCGCGGTGCAATC 420
Db 4450 TGGTGTGTGTCGATGAGACGAGCGGCTGAAGCCTGTAAGCGCGGTGCAATC 4391
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCGCGCATCAACTGGGTGCCAGCG 480
Db 4390 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCGCGCATCAACTGGGTGCCAGCG 4331
QY 481 TTGCTGTGAGAGTGGCTGCACTAATGTTCCGGCGTATTTCTGATGCTCTGACCCAGA 540
Db 4330 TTGCTGTGAGAGTGGCTGCACTAATGTTCCGGCGTATTTCTGATGCTCTGACCCAGA 4271
QY 541 CACCCATCAACAGTATTTCTCCCATGAAGAGCGGTACGCACTGGCGGTGGAGATC 600
Db 4270 CACCCATCAACAGTATTTCTCCCATGAAGAGCGGTACGCACTGGCGGTGGAGATC 4211
QY 601 TGGTGCATGGGTGCACGAGAAATGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Db 4210 TGGTGCATGGGTGCACGAGAAATGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG 4151
QY 661 CGCGTCTGGCTGGCTGGCTGCAATAATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db 4150 CGCGTCTGGCTGGCTGGCTGCAATAATCTCACTCGCAATCAAAATTCAGCGGATAG 4091
QY 721 CGGAAACGGGAAGCGACATGGAGTGCATGTCCGGTTCCTCAACAAACCATGCAATGCTGA 780
Db 4090 CGGAAACGGGAAGCGACATGGAGTGCATGTCCGGTTCCTCAACAAACCATGCAATGCTGA 4031
QY 781 ATGAGGGCATGTTCCCACTGGCATGCTGGTGGCGCAACGATCAGATGGCGGTGGCGCAA 840
Db 4030 ATGAGGGCATGTTCCCACTGGCATGCTGGTGGCGCAACGATCAGATGGCGGTGGCGCAA 3971
QY 841 TCGCGCCATACCGAGTCCGGCTGCGCTGCTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db 3970 TCGCGCCATACCGAGTCCGGCTGCGCTGCTGGTGGCGATATCTCGGTAGTGGGATACG 3911
QY 901 ACGATACCGAAGACAGCTCATGTTATTCGCGCGTTAAACCAATCAACAGGATTTTC 960
Db 3910 ACGATACCGAAGACAGCTCATGTTATTCGCGCGTTAAACCAATCAACAGGATTTTC 3851
QY 961 GCCTGCTGGGCAACACAGCTGGACCGCTTGTGCAACTCTCTCAGGCGCAGCGGTGA 1020
Db 3850 GCCTGCTGGGCAACACAGCTGGACCGCTTGTGCAACTCTCTCAGGCGCAGCGGTGA 3791
QY 1021 AGGGCAATCAGCTGTGCGCTCTCACTGGTGAAGAAAAAACACCCCTGGCGCCCAATA 1080
Db 3790 AGGGCAATCAGCTGTGCGCTCTCACTGGTGAAGAAAAAACACCCCTGGCGCCCAATA 3731
QY 1081 CGCAACCGCGCTCTCCCGCGCGTGGCCGATTCATTAATGCACTGGCAGCAGAGTTT 1140

Db 3730 CCGAACCCGCTCTCCCGCGCTTGCCTGATTCATTAATGAGCTGGCAGCAGAGTTT 3671
QY 1141 CCCGACTGAAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTGTAGCTCACTCAATAG 1200
Db 3670 CCCGACTGAAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTGTAGCTCACTCAATAG 3611
QY 1201 GCAC 1204
Db 3610 GCAC 3607
RESULT 13
US-09-702-705-785/c
; Sequence 785, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Badgur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-785
Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATATAGCGCCCGGAAGAGA 60
Db 4677 CGGACACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATATAGCGCCCGGAAGAGA 4618
QY 61 GTCATTTAGGGTGGTGAATGTGAACACAGTAAAGTTATACGATGTCGAGATATGCCG 120
Db 4617 GTCATTTAGGGTGGTGAATGTGAACACAGTAAAGTTATACGATGTCGAGATATGCCG 4558
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACAGGCGCAGCTTTCTGCGAAAA 180
Db 4557 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACAGGCGCAGCTTTCTGCGAAAA 4498
QY 181 CGCGGGAAGAGTGAAGCGCGCATGGCGAGCTGAATTCATTCCTCAACCGCGTGGCAC 240
Db 4497 CGCGGGAAGAGTGAAGCGCGCATGGCGAGCTGAATTCATTCCTCAACCGCGTGGCAC 4438
QY 241 AACAACTGGCGGGCAACAGCTGCTGTGATGTCGCGGTGGCGATTCCTCAGCTTGGCCCTGC 300
Db 4437 AACAACTGGCGGGCAACAGCTGCTGTGATGTCGCGGTGGCGATTCCTCAGCTTGGCCCTGC 4378
QY 301 ACGGCGCGTGGCAAAATGTCGCGCGGATTAATCTCGCGCGCATCAACTGGGTGCCAGCG 360
Db 4377 ACGGCGCGTGGCAAAATGTCGCGCGGATTAATCTCGCGCGCATCAACTGGGTGCCAGCG 4318
QY 361 TGGTGTGTGTCGATGAGACGAGCGCGTGAAGCCTGTAAAGCGCGGTGCAATC 420
Db 4317 TGGTGTGTGTCGATGAGACGAGCGCGTGAAGCCTGTAAAGCGCGGTGCAATC 4258
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTCATTAATGCACTGGCAGCAGAGTTT 480

```
Db 4257 TTCTCGCGCAACGGCTCAGTGGGCTGATCATTAATCTATCCGTGATGACCAAGGATGCCA 4198
Qy 481 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCCGGCGTATTCTTCTGATGTCCTGACCCAGA 540
Db 4197 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCCGGCGTATTCTTCTGATGTCCTGACCCAGA 4138
Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGCGTGGAGCATC 600
Db 4137 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGCGTGGAGCATC 4078
Qy 601 TGGTGCATTTGGTTCACAGCAAAATCGGCTGTTAGCGGGCCCAATTAAGTTCGTCTCGG 660
Db 4077 TGGTGCATTTGGTTCACAGCAAAATCGGCTGTTAGCGGGCCCAATTAAGTTCGTCTCGG 4018
Qy 661 CGCGTCTGCGTCTGGCTGGCATATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4017 CGCGTCTGCGTCTGGCTGGCATATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 3958
Qy 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTCCTCAACAAACCATGCAAAATGCTGA 780
Db 3957 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTCCTCAACAAACCATGCAAAATGCTGA 3898
Qy 781 ATGAGGGCATCTTCCCACTGCGATGCTGGTTCGCAACGATGCGCTGGGCGCA 840
Db 3897 ATGAGGGCATCTTCCCACTGCGATGCTGGTTCGCAACGATGCGCTGGGCGCA 3838
Qy 841 TCGCGCCATTAACGAGTCCGGCTCGCGTTCGGTTCGGATATCTCGGTAGTGGGATACG 900
Db 3837 TCGCGCCATTAACGAGTCCGGCTCGCGTTCGGTTCGGATATCTCGGTAGTGGGATACG 3778
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACCAACCATCAACAGGATTTTC 960
Db 3777 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACCAACCATCAACAGGATTTTC 3718
Qy 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Db 3717 GCCTGCTGGGCAACACAGCGTGGACCGCTTCTGCAACTCTCTCAGGCGCCAGGCGGTGA 3658
Qy 1021 AGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCCAATA 1080
Db 3657 AGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCCAATA 3598
Qy 1081 CGCAACCGCCTCTCCCGCGGCTTGGCGGATTCATTATGAGCTGGCACACAGGTTT 1140
Db 3597 CGCAACCGCCTCTCCCGCGGCTTGGCGGATTCATTATGAGCTGGCACACAGGTTT 3538
Qy 1141 CCGGACTGGAAGCGGCGATGAGCGCAACGCAATTAATGTCAGTTAGTCACTCATTTAG 1200
Db 3537 CCGGACTGGAAGCGGCGATGAGCGCAACGCAATTAATGTCAGTTAGTCACTCATTTAG 3478
Qy 1201 GCAC 1204
Db 3477 GCAC 3474
```

RESULT 14

```
US-09-736-457-785/c
; Sequence 785, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
```

```
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-785

Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CGGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGGCATGATAGCCCGCGGAAGAGA 60
Db 4677 CGGACACCATCGAATGGCGCAAAACCTTTTCGGGTATGGCATGATAGCCCGCGGAAGAGA 4618
Qy 61 GTCAATTCAGGGTGGTGAATGTAACCAAGTAAAGTTATACGATGTCGAGAGTATGCG 120
Db 4617 GTCAATTCAGGGTGGTGAATGTAACCAAGTAAAGTTATACGATGTCGAGAGTATGCG 4558
Qy 121 GTGTCCTTTATCAGACCGCTTTCCCGCGTGTGAACCAAGCCAGCCACGCTTTCTGCAAAA 180
Db 4557 GTGTCCTTTATCAGACCGCTTTCCCGCGTGTGAACCAAGCCAGCCACGCTTTCTGCAAAA 4498
Qy 181 CGCGGAAAAAGTGAAGCGCGCATGGCGAGCTGAATTACATTTCCCAACCGCGTGGCAC 240
Db 4497 CGCGGAAAAAGTGAAGCGCGCATGGCGAGCTGAATTACATTTCCCAACCGCGTGGCAC 4438
Qy 241 AACAACTGGCGGGCAACAGTCGTTGCTGANTGGGTTGCCACCTCCAGTCTGGGCCCTGC 300
Db 4437 AACAACTGGCGGGCAACAGTCGTTGCTGANTGGGTTGCCACCTCCAGTCTGGGCCCTGC 4378
Qy 301 ACGCGCGCTCGCAAAATTCGCGGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
Db 4377 ACGCGCGCTCGCAAAATTCGCGGGCGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCG 4318
Qy 361 TGGTGGTTCGATGGTGAAGCAAGCGCGCTCGAAGCTCTGAAGCGCGGTGCAAAATC 420
Db 4317 TGGTGGTTCGATGGTGAAGCAAGCGCGCTCGAAGCTCTGAAGCGCGGTGCAAAATC 4258
Qy 421 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATCTCCGTGATGACCAAGGATGCCA 480
Db 4257 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATCTCCGTGATGACCAAGGATGCCA 4198
Qy 481 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCCGGCGTATTCTTCTGATGTCCTGACCCAGA 540
Db 4197 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCCGGCGTATTCTTCTGATGTCCTGACCCAGA 4138
Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGCGCGTGGAGCATC 600
Db 4137 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGCGCGTGGAGCATC 4078
Qy 601 TGGTGCATTTGGTTCACAGCAAAATCGGCTGTTAGCGGGCCCAATTAAGTTCGTCTCGG 660
Db 4077 TGGTGCATTTGGTTCACAGCAAAATCGGCTGTTAGCGGGCCCAATTAAGTTCGTCTCGG 4018
Qy 661 CGCGTCTGCGTCTGGCTGGCATATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4017 CGCGTCTGCGTCTGGCTGGCATATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 3958
Qy 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTCCTCAACAAACCATGCAAAATGCTGA 780
Db 3957 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTCCTCAACAAACCATGCAAAATGCTGA 3898
Qy 781 ATGAGGGCATCTTCCCACTGCGATGCTGGTTCGCAACGATGCGCTGGGCGCA 840
Db 3897 ATGAGGGCATCTTCCCACTGCGATGCTGGTTCGCAACGATGCGCTGGGCGCA 3838
Qy 841 TCGCGCCATTAACGAGTCCGGCTCGCGTTCGGTTCGGATATCTCGGTAGTGGGATACG 900
Db 3837 TCGCGCCATTAACGAGTCCGGCTCGCGTTCGGTTCGGATATCTCGGTAGTGGGATACG 3778
```

```
QY 901 ACGATACCGAAGACAGCTCATGTATATCCCGCGTTAACCAACCATCAAAACAGGATTTTC 960
Db 3777 ACGATACCGAAGACAGCTCATGTATATCCCGCGTTAACCAACCATCAAAACAGGATTTTC 3718
QY 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTCTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 3717 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTCTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 3658
QY 1021 AGGGCAATCAGCTGTGCGCGCTCTCACTGTGTAAGAAAGAAACCAACCTTGGCGGCCAATA 1080
Db 3657 AGGGCAATCAGCTGTGCGCGCTCTCACTGTGTAAGAAAGAAACCAACCTTGGCGGCCAATA 3598
QY 1081 CGCAACCGCTCTCCCGCGCTTGGCGGATTCAATTAATGAGCTGGCAGACAGGTTT 1140
Db 3597 CGCAACCGCTCTCCCGCGCTTGGCGGATTCAATTAATGAGCTGGCAGACAGGTTT 3538
QY 1141 CCCGACTGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTAGCTCACTCATTTAG 1200
Db 3537 CCCGACTGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTAGCTCACTCATTTAG 3478
QY 1201 GCAC 1204
Db 3477 GCAC 3474
```

RESULT 15

```
US-09-614-124B-785/c
; Sequence 785, Application US/09614124B
; Patent No. 630574
```

GENERAL INFORMATION:

```
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Derrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
```

```
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
```

```
; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
```

```
; CURRENT APPLICATION NUMBER: US/09/614,124B
```

```
; CURRENT FILING DATE: 2001-07-11
```

```
; NUMBER OF SEQ ID NOS: 1668
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
```

```
; SEQ ID NO 785
```

```
; LENGTH: 5502
```

```
; TYPE: DNA
```

```
; ORGANISM: Homo sapien
```

```
US-09-614-124B-785
```

```
Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGTGCAAAACCTTCCGGTATGGCATGATAGCGCCGGAAGAGA 60
Db 4677 CGGACACCATCGAATGTGCGCAAAACCTTCCGGTATGGCATGATAGCGCCGGAAGAGA 4618
QY 61 GTCAATTTCAGGTGGTGAATGTGAACAGTAACGTTATACGATGTGCGAGATATCCG 120
Db 4617 GTCAATTTCAGGTGGTGAATGTGAACAGTAACGTTATACGATGTGCGAGATATCCG 4558
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCTGGTGAACACGAGCCAGCTTCTGCGAAA 180
Db 4557 GTGTCTCTTATCAGACCGTTTCCCGCTGGTGAACACGAGCCAGCTTCTGCGAAA 4498
QY 181 CGCGGGAAAAAGTGGAGCGCGCATGCGGAGCTGAATTACATTCCCAACCGGTGGCAC 240
Db 4497 CGCGGGAAAAAGTGGAGCGCGCATGCGGAGCTGAATTACATTCCCAACCGGTGGCAC 4438
QY 241 AACCACTGGCGGGCAACAGTCGTGTGCTGATTGGCGTTGCCACTCCAGCTCGGCCCTGC 300
```

```
Search completed: August 3, 2005, 02:32:18
Job time : 357.4 sec
```

```
Db 4437 AACCACTGGCGGGCAACAGTCGTGTGCTGATTGGCGTTGCCACTCCAGTCTGGCCCTGC 4378
QY 301 ACGCGCGTTCGCAAAATTTGTCGCGCGATTAAATCTCGCGCGCATCAACTGGGTCCGACGC 360
Db 4377 ACGCGCGTTCGCAAAATTTGTCGCGCGATTAAATCTCGCGCGCATCAACTGGGTCCGACGC 4318
QY 361 TGTGTGTGTGATGTTAGAAACGAAGCGCGTTCGAAGCTCTGTAAGCGCGGTGCACAATC 420
Db 4317 TGTGTGTGTGATGTTAGAAACGAAGCGCGTTCGAAGCTCTGTAAGCGCGGTGCACAATC 4258
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAACCTATCCGCTGGATGACAGAGATGCCA 480
Db 4257 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAACCTATCCGCTGGATGACAGAGATGCCA 4198
QY 481 TTGCTGTGGAAGCTGCTGCACATAATTTCCGGCGTTATTTCTTTGATGTCTCTGACCAGA 540
Db 4197 TTGCTGTGGAAGCTGCTGCACATAATTTCCGGCGTTATTTCTTTGATGTCTCTGACCAGA 4138
QY 541 CACCATCAACAGTATTTATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 600
Db 4137 CACCATCAACAGTATTTATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 4078
QY 601 TGTTCGCATTGGGTACACCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db 4077 TGTTCGCATTGGGTACACCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 4018
QY 661 CGGCTGTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4017 CGGCTGTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 3958
QY 721 CGGAACGGGAAGCGGACTGGAGTGCCATGTCGGTTTTTCAACAAACCATGCAATGCTGA 780
Db 3957 CGGAACGGGAAGCGGACTGGAGTGCCATGTCGGTTTTTCAACAAACCATGCAATGCTGA 3898
QY 781 ATGAGGCGCATCGTTCCCACTGCGATGCTGTTGCCAAACGATCAGATGGCGCTGGCGCAA 840
Db 3897 ATGAGGCGCATCGTTCCCACTGCGATGCTGTTGCCAAACGATCAGATGGCGCTGGCGCAA 3838
QY 841 TGGCGCGCATTCAGGATCGGGCTCGGCTGGTGGATATCTCGGTAGTGGGATACG 900
Db 3837 TGGCGCGCATTCAGGATCGGGCTCGGCTGGTGGATATCTCGGTAGTGGGATACG 3778
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAAAACAGGATTTTC 960
Db 3777 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAAAACAGGATTTTC 3718
QY 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 3717 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 3658
QY 1021 AGGGCAATCAGCTGTGTTGGCGCTCTCACTGGTGAAGAAAGAAACCAACCTTGGCGGCCAATA 1080
Db 3657 AGGGCAATCAGCTGTGTTGGCGCTCTCACTGGTGAAGAAAGAAACCAACCTTGGCGGCCAATA 3598
QY 1081 CGCAACCGCTCTCCCGCGCGTTGGCGGATTCAATTAATGAGCTGGCAGCTGGACAGAGTTT 1140
Db 3597 CGCAACCGCTCTCCCGCGCGTTGGCGGATTCAATTAATGAGCTGGCAGCTGGACAGAGTTT 3538
QY 1141 CCCGACTGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Db 3537 CCCGACTGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 3478
QY 1201 GCAC 1204
Db 3477 GCAC 3474
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 20:22:50 ; Search time 9276 Seconds
(without alignments)
17238.272 Million cell

Title: US-09-765-555B-18

Perfect score: 3300
Sequence: 1 ccgacaccatcgaaatggtgc.....acgacgttccggactacgct 3300

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

100% Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: ★

1: qb ba:★

2: qb_htq: *

3: gb_in:*

4: $ab_{-}om: \star$

5. $\frac{1}{2} - \frac{1}{3} = \frac{1}{6}$ ✓

6: qb_nat.*

6: gb_pat:*

9: gb_pn:*

ॐ नमो भगवते वासुदेवाय

9: gb_pr: *

10: gb_ro: *

```
11: gb_stg:
```

12: gb_sy: *

13: gb_un: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES									
Result No.	Score	Query Match	Length	DB ID	Description				
1	3300	100.0	3300	6	AX202430	AX202430	Sequence		
2	3322.8	98.0	3300	6	AX202429	AX202429	Sequence		
3	3226.4	97.8	3300	6	AX202426	AX202426	Sequence		
4	3226.4	97.8	3300	6	AX202427	AX202427	Sequence		
5	3216.8	97.5	3300	6	AX202428	AX202428	Sequence		
6	2705.8	82.0	7373	6	AX284144	AX284144	Sequence		
7	2705.8	82.0	8101	6	AX172306	AX172306	Sequence		
8	2701	81.8	6648	6	AX377531	AX377531	Sequence		
9	2700	81.8	7475	6	A91365	A91365	Sequence 1		
10	2700	81.8	7475	6	AR031992	AR031992	Sequence		
11	2700	81.8	7475	6	AR027294	AR027294	Sequence		
12	2698	81.8	9191	6	AX377532	AX377532	Sequence		
13	2676.2	81.1	6724	12	AF097412	AF097412	Express		
14	2636.8	79.9	6806	6	AX378208	AX378208	Sequence		
15	2631	79.7	6706	12	AF031088	AF031088	Shuttle		
16	2598.2	78.7	6748	12	AF031813	AF031813	Express		
17	1609.4	48.8	5558	6	CQ794769	CQ794769	Sequence		
18	1474	44.7	4700	6	CQ846980	CQ846980	Sequence		
19	1462.4	44.3	4921	6	AX768201	AX768201	Sequence		

ALIGNMENTS

RESULT	1
AX202430	
LOCUS	3300 bp DNA linear PAT 30-AUG-2001
DEFINITION	Sequence 18 from Patent WO0152620.
ACCESSION	AX202430
VERSION	AX202430.1 GI:15392178
KEYWORDS	.
SOURCE	synthetic construct
ORGANISM	synthetic construct
	other sequences; artificial sequences.
REFERENCE	1
AUTHORS	Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE	Methods and compositions to modulate expression in plants
JOURNAL	Patent: WO 0152620-A 18 26-JUL-2001;
	The Scripps Research Institute (US) ; SYNGENTA AGRICULTURAL
	DISCOVERY, INC. (CA)
FEATURES	Location/Qualifiers
source	1..3300
	/organism="synthetic construct"
	/mol_type="unassigned DNA"
	/db_xref="taxon:32630"
	/note="Parial sequence of pMal-Ap3 and zinc finger protein ZFPAP3"
ORIGIN	
Query Match	100.0%; Score 3300; DB 6; Length 3300;
Best Local Similarity	100.0%; Pred.No. 0;
Matches 3300; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Qy	1 CCGACACCATCGAATGGTGCAAACCTTTTCGCGGTATGCCATGATAGCCGCCGGAAGAGA 60
Dd	1 CCGACACCATCGAATGGTGCAAACCTTTTCGCGGTATGCCATGATAGCCGCCGGAAGAGA 60
Qy	61 GTCAAATTCAAGGTGGTGAAATGTGAACCAGTAACGTTTATACGATGTCGCAGAGTATCCCG 120
Dd	61 GTCAAATTCAAGGTGGTGAAATGTGAACCAGTAACGTTTATACGATGTCGCAGAGTATCCCG 120
Qy	121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGACACGAGCCAGCCAGTTTCTCGGAAA 180
Dd	121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGACACGAGCCAGCCAGTTTCTCGGAAA 180
Qy	181 CGCGGGAAAAAGTGAAGCGCGGATGCGGAGCTGAATTATCTCCCAACCGCGTGGCAC 240

Db 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAAATTAATTCCTCCCAACCGCGTGGCAC 240
Qy 241 AACAACTGGCGGGCAAAACAGTCGTGTCTGATTTGGCGTTGGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGCAAAACAGTCGTGTCTGATTTGGCGTTGGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 AGCGCGCTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db 301 AGCGCGCTCGCAAAATTTGTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Qy 361 TGTGTGTTCATGGTAGAAGCGCGCTGAGCTGATTAATTAATTCCTGATGCGGATGCGCA 420
Db 361 TGTGTGTTCATGGTAGAAGCGCGCTGAGCTGATTAATTAATTCCTGATGCGGATGCGCA 420
Qy 421 TTTCTCGGCAACGGCTCAGTGGGCTGATCAATTAATTCCTGATGCGGATGCGCA 480
Db 421 TTTCTCGGCAACGGCTCAGTGGGCTGATCAATTAATTCCTGATGCGGATGCGCA 480
Qy 481 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTTAATTTCTTGATGTTCTCTGACCAGA 540
Db 481 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTTAATTTCTTGATGTTCTCTGACCAGA 540
Qy 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACCGGACTGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACCGGACTGGCGTGGAGCATC 600
Qy 601 TGGTCGATTTGGGTACCCAGCAATTCGCGTGTAGCGGGCCCATTAAGTTCTGCTCGG 660
Db 601 TGGTCGATTTGGGTACCCAGCAATTCGCGTGTAGCGGGCCCATTAAGTTCTGCTCGG 660
Qy 661 CGCGTCTGCGTCTGGCTGGCATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGAGCTGGAGTCCATGTCGGGTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGAGCTGGAGTCCATGTCGGGTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATCGTTCACACTCGCATGCTGTGTGCAACAGATCAGATGGCGTGGCGGCA 840
Db 781 ATGAGGGCATCGTTCACACTCGCATGCTGTGTGCAACAGATCAGATGGCGTGGCGGCA 840
Qy 841 TGC CGGCAATTAACGAGTCCGGGCTGGCGTGTGGATATCTCGGTAGTGGGATACG 900
Db 841 TGC CGGCAATTAACGAGTCCGGGCTGGCGTGTGGATATCTCGGTAGTGGGATACG 900
Qy 901 ACGATACGGAAGACAGCTCATGTTATATCCGCGGTTAAACCAATCAAAACAGATTTTC 960
Db 901 ACGATACGGAAGACAGCTCATGTTATATCCGCGGTTAAACCAATCAAAACAGATTTTC 960
Qy 961 GCCTGTGGGCAAAACAGCGTGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGTGGGCAAAACAGCGTGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTGGCGTCTCACTGTGGAAGAAACCAACCTGGCGGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGGCGTCTCACTGTGGAAGAAACCAACCTGGCGGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGATTCATTAATGCGAGCTGGCAGCAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGATTCATTAATGCGAGCTGGCAGCAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG 1200
Qy 1201 GCACAAATCTCATGTTTGAACGCTTATCATGACTGCAAGGTGCAACAAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGAACGCTTATCATGACTGCAAGGTGCAACAAATGCTTCTGGCG 1260
Qy 1261 TCAGGAGCCCATCGGAAGCTGTGTATGGCTGTGAGGTGTAATCACTGCATAATTCG 1320

Db 1261 TCAGGAGCCATCGGAAGCTGTGTATGGCTGTGAGGTGTAATCACTGCATAATTCG 1320
Qy 1321 TGTGCTCAAGGGGCACTCCCGTTCTCGATAATGTTTTTTTGGCCGACATCATACCGGTT 1380
Db 1321 TGTGCTCAAGGGGCACTCCCGTTCTCGATAATGTTTTTTTGGCCGACATCATACCGGTT 1380
Qy 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA 1440
Qy 1441 ATTTGTAGCGGATCAAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCAGGA 1500
Db 1441 ATTTGTAGCGGATCAAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTTCAGGA 1500
Qy 1501 GCATCTCAACAAAGGACCATGATTAATGAAGCTGAAAGCTGAAAGTAACTGGTAAATCTGG 1560
Db 1501 GCATCTCAACAAAGGACCATGATTAATGAAGCTGAAAGCTGAAAGTAACTGGTAAATCTGG 1560
Qy 1561 ATTAACCGGATAAAGGCTATAACGGTCTCGCTGAACTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACCGGATAAAGGCTATAACGGTCTCGCTGAACTCGGTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCGCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCGCCACAGGTT 1680
Qy 1681 GCGGCAACTGGCGATGGCGCTGACATTAATCTCTGGGCAACAGCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCGCTGACATTAATCTCTGGGCAACAGCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCGTGTGGCTGAAATCAACCGGCAAAAGCTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCGTGTGGCTGAAATCAACCGGCAAAAGCTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTCAACCGGCAAGCTGATGCTTACCGATGCGTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTCAACCGGCAAGCTGATGCTTACCGATGCGTGT 1860
Qy 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCGCCCAAAACCTCGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCGCCCAAAACCTCGGAA 1920
Qy 1921 GAGATCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGTATGTTCAAC 1980
Db 1921 GAGATCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGTATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGGTACTTCACTGCGCGCTGATTTGCTGCTGACCGGGGTTATGCTTCAAG 2040
Db 1981 CTGCAAGAACCGGTACTTCACTGCGCGCTGATTTGCTGCTGACCGGGGTTATGCTTCAAG 2040
Qy 2041 TATGAAACCGGCAAGTACGACATTAAGACGTGGCGTGGATAAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGGCAAGTACGACATTAAGACGTGGCGTGGATAAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGTTGATGACCTGATTAATAAACAACATGAATGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTGATGACCTGATTAATAAACAACATGAATGAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTTAAAGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAAGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGTTCCAACTGACACAGCAAGTGAATTAATGTTGTAACGGTACTTCCGACCTTC 2280
Db 2221 GCATGTTCCAACTGACACAGCAAGTGAATTAATGTTGTAACGGTACTTCCGACCTTC 2280
Qy 2281 AAGGTCACCACTCCAAAACCGTTCTGTTGGCGTCTGAGCGCAGGTATTAAACGGCCAGT 2340
Db 2281 AAGGTCACCACTCCAAAACCGTTCTGTTGGCGTCTGAGCGCAGGTATTAAACGGCCAGT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCTGAAACTATCTGCTGACTGATGAGGTTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCTGAAACTATCTGCTGACTGATGAGGTTCTG 2400

Qy	2401	GAAGCGGTAAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAG	2460
Db	2401	GAAGCGGTAAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAG	2460
Qy	2461	TTGGCGAAGATCCACGCTATTGGCGCCACCATGGAAAAAGCCGACAGAAAGGTGAATCATG	2520
Db	2461	TTGGCGAAGATCCACGCTATTGGCGCCACCATGGAAAAAGCCGACAGAAAGGTGAATCATG	2520
Qy	2521	CCGAACATCCCGCAGATGTCGCTTTCTTGATGTCGCTGCGTACTGCGGTGATCAACGCC	2580
Db	2521	CCGAACATCCCGCAGATGTCGCTTTCTTGATGTCGCTGCGTACTGCGGTGATCAACGCC	2580
Qy	2581	GCCAGCGTGTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG	2640
Db	2581	GCCAGCGTGTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG	2640
Qy	2641	AACAAACAACAATAACAATAACAACACCTCGGGATCGAGGGAAGGATTTTCAAAATTC	2700
Db	2641	AACAAACAACAATAACAATAACAACACCTCGGGATCGAGGGAAGGATTTTCAAAATTC	2700
Qy	2701	GGATCCTCTTCTGTCGCGCCAGCGGCCCTTCGAGCCCGGGGAGAACCTTATGCTTGT	2760
Db	2701	GGATCCTCTTCTGTCGCGCCAGCGGCCCTTCGAGCCCGGGGAGAACCTTATGCTTGT	2760
Qy	2761	CCGGAATGTGTAAGTCTTTCAGCCAGAGCAGTCCCTGGTGGCCACACAGCGTACCCAC	2820
Db	2761	CCGGAATGTGTAAGTCTTTCAGCCAGAGCAGTCCCTGGTGGCCACACAGCGTACCCAC	2820
Qy	2821	ACGGGTGAAGAACCGTATTAATGCCAGAGTGGCGCAATCTTTTAGCAGTCCAGCAAC	2880
Db	2821	ACGGGTGAAGAACCGTATTAATGCCAGAGTGGCGCAATCTTTTAGCAGTCCAGCAAC	2880
Qy	2881	CTGGTGGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCCAGAAATGTGGC	2940
Db	2881	CTGGTGGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCCAGAAATGTGGC	2940
Qy	2941	AAGTCTTTTCAGCCAGTCCAGCAACTGTGTGGCCACCAACGTACTCACACGGGGAGAAG	3000
Db	2941	AAGTCTTTTCAGCCAGTCCAGCAACTGTGTGGCCACCAACGTACTCACACGGGGAGAAG	3000
Qy	3001	CCCTATGCTTGCCTGGATGTGTAAGTCCCTTTCAGCACGAGTGGCTCTTGGTTAGACAC	3060
Db	3001	CCCTATGCTTGCCTGGATGTGTAAGTCCCTTTCAGCACGAGTGGCTCTTGGTTAGACAC	3060
Qy	3061	CAGCGTACCCACACGGGTGAAAAACGTTAATATGCCAGAGTGGCGGAAATCTTTTAGC	3120
Db	3061	CAGCGTACCCACACGGGTGAAAAACGTTAATATGCCAGAGTGGCGGAAATCTTTTAGC	3120
Qy	3121	CAGCGCGCCACCTGGAAAGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT	3180
Db	3121	CAGCGCGCCACCTGGAAAGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT	3180
Qy	3181	CCAGAAATGTGGCAAGTCTTCTCAACTTCAGCAACTGGTCCGTCACCAAGTACTAC	3240
Db	3181	CCAGAAATGTGGCAAGTCTTCTCAACTTCAGCAACTGGTCCGTCACCAAGTACTAC	3240
Qy	3241	ACCGGTAAAAAATAGTGGCCAGCGCCGACAGTACCGGTACGAGTTCGCGACTACGCT	3300
Db	3241	ACCGGTAAAAAATAGTGGCCAGCGCCGACAGTACCGGTACGAGTTCGCGACTACGCT	3300
RESULT 2	AX202429	Sequence 17 from Patent WO0152620.	
LOCUS	AX202429	3300 bp	DNA
DEFINITION	AX202429	linear	PAT 30-AUG-2001
ACCESSION	AX202429		
VERSION	AX202429.1		
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE	1		

AUTHORS	Barbas, C.F., Stege, J.T., Guan, X. and Dalmia, B.									
TITLE	Methods and compositions to modulate expression in plants									
JOURNAL	Patent: WO 0152620-A 17 26-JUL-2001;									
	The Scripps Research Institute (US) ; SYNGENTA AGRICULTURAL									
	DISCOVERY, INC. (CA)									
FEATURES	Location/Qualifiers									
source	1..3300									
	/organism="synthetic construct"									
	/mol_type="unassigned DNA"									
	/db_xref="taxon:32630"									
	/note="Partial sequence of pMal-m4 and zinc finger protein zFPm4"									
ORIGIN										
Query Match	98.0%; Score 3232.8; DB 6; Length 3300;									
Best Local Similarity	98.7%; Pred. No. 0;									
Matches 3258; Conservative	0; Mismatches 42; Indels 0; Gaps 0;									
QY	1	CCGACACCATCGAATGGTGC	AAACCTTTTCGCGGTATG	CGCATGATAGCGCCCGGA	GAGA	60				
Db	1	CCGACACCATCGAATGGTGC	AAACCTTTTCGCGGTATG	CGCATGATAGCGCCCGGA	GAGA	60				
QY	61	GTCAATTGAGGTGGTGAATG	TGAACACCACTAACGTTAT	ACGATGTCGACAGATATG	CCG	120				
Db	61	GTCAATTGAGGTGGTGAATG	TGAACACCACTAACGTTAT	ACGATGTCGACAGATATG	CCG	120				
QY	121	GTGTCTCTTATCAGACCG	TTTCCC	CGGTGGTGAACACGAGCCAG	CCACGTTTCTGCG	180				
Db	121	GTGTCTCTTATCAGACCG	TTTCCC	CGGTGGTGAACACGAGCCAG	CCACGTTTCTGCG	180				
QY	181	CGCGGAAAAAGTGAAGCG	GGGATGGCGGAGCTGAA	TTTACATTCCTCCACCGCGT	GCCAC	240				
Db	181	CGCGGAAAAAGTGAAGCG	GGGATGGCGGAGCTGAA	TTTACATTCCTCCACCGCGT	GCCAC	240				
QY	241	AACAACCTGCGCGGCAAC	AGTCTGTTCTGATTGGCG	GTGACCTCCAGTCTCG	CCCTGC	300				
Db	241	AACAACCTGCGCGGCAAC	AGTCTGTTCTGATTGGCG	GTGACCTCCAGTCTCG	CCCTGC	300				
QY	301	ACGCGCGCTCGCAAA	TTGTTCGCGCGATTTAA	ATCTCGCGCGATCAACTGG	GTGCCAGCG	360				
Db	301	ACGCGCGCTCGCAAA	TTGTTCGCGCGATTTAA	ATCTCGCGCGATCAACTGG	GTGCCAGCG	360				
QY	361	TGTTGTTGTCGATGTTAG	AAACGAGCGGCTCGAAG	CGCTGTAAAGCGCGGTGC	ACAATC	420				
Db	361	TGTTGTTGTCGATGTTAG	AAACGAGCGGCTCGAAG	CGCTGTAAAGCGCGGTGC	ACAATC	420				
QY	421	TTCTCGCGCAACGGTCA	GTGGGTGATCAATTAAC	TATCCGCTGGATGAC	ACGAGATGCCA	480				
Db	421	TTCTCGCGCAACGGTCA	GTGGGTGATCAATTAAC	TATCCGCTGGATGAC	ACGAGATGCCA	480				
QY	481	TTGCTGTGGAAGCTGCT	GACATAATGTTCCGGCG	TTATTTCTTGATGTCTCT	GACCAGA	540				
Db	481	TTGCTGTGGAAGCTGCT	GACATAATGTTCCGGCG	TTATTTCTTGATGTCTCT	GACCAGA	540				
QY	541	CACCCATCAACAGTATTA	TTTTCTCCCATGAAGAC	GGTACGCACTGGCGGT	GGAGCATC	600				
Db	541	CACCCATCAACAGTATTA	TTTTCTCCCATGAAGAC	GGTACGCACTGGCGGT	GGAGCATC	600				
QY	601	TGTTCCGATTTGGGTCA	CCAGCAAAATCGCGT	GTTAGCGGCGCCATTAAG	TTTCTGTCTCGG	660				
Db	601	TGTTCCGATTTGGGTCA	CCAGCAAAATCGCGT	GTTAGCGGCGCCATTAAG	TTTCTGTCTCGG	660				
QY	661	CGCGTCTGCGTCTGG	CTGGCATAAATATCTCA	CTCGCAATCAAAATTCAG	CCGATAG	720				
Db	661	CGCGTCTGCGTCTGG	CTGGCATAAATATCTCA	CTCGCAATCAAAATTCAG	CCGATAG	720				
QY	721	CGGAACGGGAAGCG	AGCTGGAGTGC	ATGTTTCAACAAACCATG	CAAAATGCTGA	780				
Db	721	CGGAACGGGAAGCG	AGCTGGAGTGC	ATGTTTCAACAAACCATG	CAAAATGCTGA	780				
QY	781	ATGAGGCGATCGT	TTCCCACTGCGAT	GCTGGTTCGCAACGAT	TCAGATGGCGCTGGG	CGCA	840			
Db	781	ATGAGGCGATCGT	TTCCCACTGCGAT	GCTGGTTCGCAACGAT	TCAGATGGCGCTGGG	CGCA	840			

QY	841	TGCGGCCATTACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGGCCATTACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAACAGGATTTTC	960
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAACAGGATTTTC	960
QY	961	GCCTGCTGGGCAACACCGTGGACCGCTTGTCTGCAACTCTCTAGGCCACGCGGTGA	1020
Db	961	GCCTGCTGGGCAACACCGTGGACCGCTTGTCTGCAACTCTCTAGGCCACGCGGTGA	1020
QY	1021	AGGCAATCAGCTGTTGCGCGCTCTCACTGGTGAAAGAAACCAACCCCTGGCGCCCAATA	1080
Db	1021	AGGCAATCAGCTGTTGCGCGCTCTCACTGGTGAAAGAAACCAACCCCTGGCGCCCAATA	1080
QY	1081	CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGAGTGTAGCTCACTCATTTAG	1140
Db	1081	CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGAGTGTAGCTCACTCATTTAG	1140
QY	1141	CCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTGTAGCTCACTCATTTAG	1200
Db	1141	CCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTGTAGCTCACTCATTTAG	1200
QY	1201	GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGACGCGTGACCAATGCTTCTGGCG	1260
Db	1201	GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGACGCGTGACCAATGCTTCTGGCG	1260
QY	1261	TCAGGACGCAATCGAAGCTGTGATGCTGTGCGAGTGTGAGTGTAGCTCACTCATTTAG	1320
Db	1261	TCAGGACGCAATCGAAGCTGTGATGCTGTGCGAGTGTGAGTGTAGCTCACTCATTTAG	1320
QY	1321	TGTCGCTCAAGCGGCACCTCCCGTTCTGGATAATGTTTTTCGCGCGACATCAACCGTT	1380
Db	1321	TGTCGCTCAAGCGGCACCTCCCGTTCTGGATAATGTTTTTCGCGCGACATCAACCGTT	1380
QY	1381	CTGGCAATATCTGAAATGAGTGTGACAAATTAATCATCGGCTCGTATAATGTGTGA	1440
Db	1381	CTGGCAATATCTGAAATGAGTGTGACAAATTAATCATCGGCTCGTATAATGTGTGA	1440
QY	1441	ATTGTAGCGGATTAACAAATTTACACAGGAACAGCGAGTCCGTTAGGTGTTTTCAGGA	1500
Db	1441	ATTGTAGCGGATTAACAAATTTACACAGGAACAGCGAGTCCGTTAGGTGTTTTCAGGA	1500
QY	1501	GCACTTCAACCAAGGACCATAGATTATGAAACTGAAGAAAGTAACTGGTAATCTGG	1560
Db	1501	GCACTTCAACCAAGGACCATAGATTATGAAACTGAAGAAAGTAACTGGTAATCTGG	1560
QY	1561	ATTAAACCGGATTAAGGCTATAACCGTCTCGCTGAAGTTCGTAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAAACCGGATTAAGGCTATAACCGTCTCGCTGAAGTTCGTAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAAGTACCGTTGAGCATCCGATATACTGGAAGAGAAATTCGCCACAGTT	1680
Db	1621	ACCGGAATTAAGTACCGTTGAGCATCCGATATACTGGAAGAGAAATTCGCCACAGTT	1680
QY	1681	CGGCAACTGCGGATGCGCTGACATTTATCTTGGGCAACAGCAGCTTTGGTGGCTAC	1740
Db	1681	CGGCAACTGCGGATGCGCTGACATTTATCTTGGGCAACAGCAGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCTGTTGGCTGAATCACCCCGGACAAAGCGTTTCCAGGACAAGCTGTAT	1800
Db	1741	GCTCAATCTGGCTGTTGGCTGAATCACCCCGGACAAAGCGTTTCCAGGACAAGCTGTAT	1800
QY	1801	CCGTTTACCTGGATGCGCTACGTTTACAAAGCTGATTGCTTACCGGATCGCTGTT	1860
Db	1801	CCGTTTACCTGGATGCGCTACGTTTACAAAGCTGATTGCTTACCGGATCGCTGTT	1860
QY	1861	GAAGCGTTATCGCTGATTATTAACAAAGATCTGTCGCGAACCCGCAAAAACCTGGAA	1920
Db	1861	GAAGCGTTATCGCTGATTATTAACAAAGATCTGTCGCGAACCCGCAAAAACCTGGAA	1920

QY	1921	GAGATCCGGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCGGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAAACCGTACTTTCACCTGCGCGCTGATTCCTGCTGACGGGGTTATGCGGTTCAAG	2040
Db	1981	CTGCAAGAAACCGTACTTTCACCTGCGCGCTGATTCCTGCTGACGGGGTTATGCGGTTCAAG	2040
QY	2041	TATGAAAAACGGCAAGTACGACATTTAAAGACGTTGGCGTGGATACGCTGGCGGAAAGCG	2100
Db	2041	TATGAAAAACGGCAAGTACGACATTTAAAGACGTTGGCGTGGATACGCTGGCGGAAAGCG	2100
QY	2101	GGTCTGACCTTCTCTGGTGTGACCTGATTTAAACCAACACATGATGACAGACCGGATTC	2160
Db	2101	GGTCTGACCTTCTCTGGTGTGACCTGATTTAAACCAACACATGATGACAGACCGGATTC	2160
QY	2161	TCCATCGCAGAAGCTGCTTTTAAATAAGGCGAAACAGCGATGACCATCAACGGCGCGTGG	2220
Db	2161	TCCATCGCAGAAGCTGCTTTTAAATAAGGCGAAACAGCGATGACCATCAACGGCGCGTGG	2220
QY	2221	GCATGTCGCAACATCGACACAGCAAAAGTGAATTTATGGTGTAAACGGTACTGCGACCTTC	2280
Db	2221	GCATGTCGCAACATCGACACAGCAAAAGTGAATTTATGGTGTAAACGGTACTGCGACCTTC	2280
QY	2281	AAGGTCACCAACATCGACACAGCAAAAGTGAATTTATGGTGTAAACGGTACTGCGACCTTC	2340
Db	2281	AAGGTCACCAACATCGACACAGCAAAAGTGAATTTATGGTGTAAACGGTACTGCGACCTTC	2340
QY	2341	CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG	2400
Db	2341	CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG	2400
QY	2401	GAAGCGGTTAATAAGCAAAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	2460
Db	2401	GAAGCGGTTAATAAGCAAAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	2460
QY	2461	TTGGCGAAAGATCCAGTATTGCGCGCACCATCGAAAAACCGCCAGAAAGGTGAATCATG	2520
Db	2461	TTGGCGAAAGATCCAGTATTGCGCGCACCATCGAAAAACCGCCAGAAAGGTGAATCATG	2520
QY	2521	CCGAAACATCCGCGAGATGTCGCGTTTCTGGTATGCGGTGCGTACTGCGGTGATCAAGCC	2580
Db	2521	CCGAAACATCCGCGAGATGTCGCGTTTCTGGTATGCGGTGCGTACTGCGGTGATCAAGCC	2580
QY	2581	GCAGCGGTCGTGAGTGTGAGTGAAGCCCTGAAAGACGCGAGACTAATTCGAGCTCG	2640
Db	2581	GCAGCGGTCGTGAGTGTGAGTGAAGCCCTGAAAGACGCGAGACTAATTCGAGCTCG	2640
QY	2641	AACAAACAAACAAATTAACAAATTAACAAACCTCGGATCGAGGAAAGGATTTTCAGAAATTC	2700
Db	2641	AACAAACAAACAAATTAACAAATTAACAAACCTCGGATCGAGGAAAGGATTTTCAGAAATTC	2700
QY	2701	GGATCTCTCTCTGTCGCGCGCGCTGAGCCCGGCGGAGAGCCCTATGCTTGT	2760
Db	2701	GGATCTCTCTCTGTCGCGCGCGCTGAGCCCGGCGGAGAGCCCTATGCTTGT	2760
QY	2761	CCGGAATGTTGAAGTCTTTCAGCAGACAGCTCCCTGCTGCGCCACAGCGTATCCAC	2820
Db	2761	CCGGAATGTTGAAGTCTTTCAGCAGACAGCTCCCTGCTGCGCCACAGCGTATCCAC	2820
QY	2821	ACGGGTGAAAAACCGTATATAATGCGGCAAACTCTTTAGCCAGTCCAGCAAC	2880
Db	2821	ACGGGTGAAAAACCGTATATAATGCGGCAAACTCTTTAGCCAGTCCAGCAAC	2880
QY	2881	CTGGTGGCGCATCAACGCACTCATCTGCGGAGAGCCCATACAAATGTCAGAAATGTCG	2940
Db	2881	CTGGTGGCGCATCAACGCACTCATCTGCGGAGAGCCCATACAAATGTCAGAAATGTCG	2940
QY	2941	AAGTCTTTTTCAGCGAGTCCAGCAACCTGGTGGCGCCACCAACGCTACTCACACCGGGGAGAG	3000
Db	2941	AAGTCTTTTTCAGCGAGTCTTTCAGCGAGTCCAGCAACCTGGTGGCGCCACCAACGCTACTCACACCGGGGAGAG	3000
QY	3001	CCCTATGCTGTGTCGGAATGTGGTAAGTCTCTTACAGCACAGTGGCTCTTGGTTAGACAC	3060

```

Db      3001  |||||
Cc      CCCTATGCTTGTCCGGAATGGTAAAGTCTCTCTCAGAGCTCTCACCTGGTGGCCAC 3060
Qy      3061  CAGCGTACCCACACGGGTGAAAAACCGGTATAAATGCCAGAGTGGCGCAAAATCTTTTACG 3120
Db      3061  CAGCGTACCCACACGGGTGAAAAACCGGTATAAATGCCAGAGTGGCGCAAAATCTTTTACG 3120
Qy      3121  CAGCGCGCCCACTTGAAGCGGCATCAAGCACTATACCTACTGGGAGAGCCATACAAATGT 3180
Db      3121  CGCAGCGATAACCTGTGGCGCATCAACGCACTCATCTGGCGAGAGCCATACAAATGT 3180
Qy      3181  CCAGAAATGGGCAAGTCTTTCTCAACTTCAGGCAACTTGGTCCGTCACCAAGTACTCAC 3240
Db      3181  CCAGAAATGGGCAAGTCTTTCTCAACTTCAGGCAACTTGGTCCGTCACCAAGTACTCAC 3240
Qy      3241  ACCGGTAAAAAATAGTGGCCAGCGCCGCGCAGTACCGGTACGAGCTTCGCGACTACGCT 3300
Db      3241  ACCGGTAAAAAATAGTGGCCAGCGCCGCGCAGTACCGGTACGAGCTTCGCGACTACGCT 3300

RESULT 3
AX202426
LOCUS      AX202426          3300 bp      DNA      linear      PAT 30-AUG-2001
DEFINITION Sequence 14 from Patent WO0152620.
ACCESSION  AX202426
VERSION     AX202426.1  GI:15392173
KEYWORDS   .
SOURCE     synthetic construct
ORGANISM   other sequences; artificial sequences.
REFERENCE  1
AUTHORS   Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE     Methods and compositions to modulate expression in plants
JOURNAL   Patent: WO 0152620-A 14 26-JUL-2001;
          The Scripps Research Institute (US) ; SYNGENTA AGRICULTURAL
          DISCOVERY, INC. (CA)

FEATURES
    source
        Location/Qualifiers
            1..3300
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
                /note="Partial sequence of pMal-m1 and zinc finger protein
                ZFPm1"

ORIGIN
    Query Match      97.8%; Score 3226.4; DB 6; Length 3300;
    Best Local Similarity 98.6%; Pred. No. 0;
    Matches 3254; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy      1  CCGACACCATCGAATGGTGC AAAACCTTTTCGGCGTATGGCATGATAGCGCCCGGAAGAGA 60
Db      1  CCGACACCATCGAATGGTGC AAAACCTTTTCGGCGTATGGCATGATAGCGCCCGGAAGAGA 60
Qy      61  GTCAAATTCAGGTGTGTAATGTGAACCAAGTAACTTATACGATGTCGCAGAGTATGCG 120
Db      61  GTCAAATTCAGGTGTGTAATGTGAACCAAGTAACTTATACGATGTCGCAGAGTATGCG 120
Qy      121  GTGTCTCTTATCAGACCGGTTTCCCGGTGGTGAACACAGGCCAGCCAGCTTTCTGCGAAAA 180
Db      121  GTGTCTCTTATCAGACCGGTTTCCCGGTGGTGAACACAGGCCAGCCAGCTTTCTGCGAAAA 180
Qy      181  CCGCGGAAAAAGTGAAGCGCGCATGGCGGAGCTGAAATTTACATTCCTCCAAACCGGTGGC 240
Db      181  CCGCGGAAAAAGTGAAGCGCGCATGGCGGAGCTGAAATTTACATTCCTCCAAACCGGTGGC 240
Qy      241  AACAACTGGCGGGCAACAGTGTGCTGATTTGGGTGGCGTGGCCACTTCAGTCTGGCCCTGC 300
Db      241  AACAACTGGCGGGCAACAGTGTGCTGATTTGGGTGGCGTGGCCACTTCAGTCTGGCCCTGC 300
Qy      301  ACGCCCGGTGCGCAAAATTTGTCGGCGGCGATTAATCTCGCGCCGATCAATCTGGGTGCCAGG 360
Db      301  ACGCCCGGTGCGCAAAATTTGTCGGCGGCGATTAATCTCGCGCCGATCAATCTGGGTGCCAGG 360
```

```

Qy      361  TGTGTTGTCGATGGTAGAACGAGCGCGTGAAGCCTGTAAAGCGCGGTGCACATC 420
Db      361  TGTGTTGTCGATGGTAGAACGAGCGCGTGAAGCCTGTAAAGCGCGGTGCACATC 420
Qy      421  TTTCTCGCGCAACGGGTCAAGTGGGCTGATCATTAATCTATCCGCTGGATGACCAAGATGCCA 480
Db      421  TTTCTCGCGCAACGGGTCAAGTGGGCTGATCATTAATCTATCCGCTGGATGACCAAGATGCCA 480
Qy      481  TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGGTTATTTCTTGATGATCTCTGACCAAGA 540
Db      481  TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGGTTATTTCTTGATGATCTCTGACCAAGA 540
Qy      541  CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACCGACTGGCGGTGGAGCATC 600
Db      541  CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACCGACTGGCGGTGGAGCATC 600
Qy      601  TGGTTCGCAATTCGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Db      601  TGGTTCGCAATTCGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Qy      661  CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db      661  CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Qy      721  CGGAACGGGAAGGCGACTGGAGTGCCATGTCGGGTTTCAACCAACCATGCAAAATGCTGA 780
Db      721  CGGAACGGGAAGGCGACTGGAGTGCCATGTCGGGTTTCAACCAACCATGCAAAATGCTGA 780
Qy      781  ATGAGGGCATCGTTCCCACTGCGATGCTGTTGTCGCAACGATCAGATGGCGTGGCGGCA 840
Db      781  ATGAGGGCATCGTTCCCACTGCGATGCTGTTGTCGCAACGATCAGATGGCGTGGCGGCA 840
Qy      841  TGGCGGCCATTTACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db      841  TGGCGGCCATTTACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
Qy      901  ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
Db      901  ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
Qy      961  GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTGTCAACTCTCTCAGGCGCAGGCGGTGA 1020
Db      961  GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTGTCAACTCTCTCAGGCGCAGGCGGTGA 1020
Qy      1021  AGGGCAATCAGCTGTGTCCTCTCACTGGTGAAGAAAAACCACTTCGCGCCCAATA 1080
Db      1021  AGGGCAATCAGCTGTGTCCTCTCACTGGTGAAGAAAAACCACTTCGCGCCCAATA 1080
Qy      1081  CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCATTAATGAGTGGTGGTGGTGGTGGTGG 1140
Db      1081  CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCATTAATGAGTGGTGGTGGTGGTGGTGG 1140
Qy      1141  CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTGGTGGTGGTGGTGGTGG 1200
Db      1141  CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTGGTGGTGGTGGTGGTGG 1200
Qy      1201  GCACAAATTCATGTTTGACAGCTTATCATCGACTGCAACGCGTGGTGGTGGTGGTGGTGG 1260
Db      1201  GCACAAATTCATGTTTGACAGCTTATCATCGACTGCAACGCGTGGTGGTGGTGGTGGTGG 1260
Qy      1261  TCAGGACGCAATCGGAAGCTGTGGTATGCTGTGAGGTGCTGAATCACTGATCAATTCG 1320
Db      1261  TCAGGACGCAATCGGAAGCTGTGGTATGCTGTGAGGTGCTGAATCACTGATCAATTCG 1320
Qy      1321  TGTGCTCAAGCGCACTCCCGTCTCGATAATGTTTTTGGCGCGAGCATCAATACGGTT 1380
Db      1321  TGTGCTCAAGCGCACTCCCGTCTCGATAATGTTTTTGGCGCGAGCATCAATACGGTT 1380
Qy      1381  CTGGCAAAATTTCTGAAATGAGCTGTGCAAAATTAATCATCGGCTCGTATATATGTTGGA 1440
Db      1381  CTGGCAAAATTTCTGAAATGAGCTGTGCAAAATTAATCATCGGCTCGTATATATGTTGGA 1440
Qy      1441  ATTGTGAGGGGATACAAATTTTCAACAGAAAAACGCCAGTCCGTTAGTGTGTTTCAAGA 1500
```

```
Db 1441 ATTGTAGCGGATAACAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTACGGA 1500
QY 1501 GCACCTTCCAAACAAAGGACCATAGATTATGAAAACTGGAAGAGGTAAATCGTAAATCTGG 1560
Db 1501 GCACCTTCCAAACAAAGGACCATAGATTATGAAAACTGGAAGAGGTAAATCGTAAATCTGG 1560
QY 1561 ATTAAACGGGATAAAGGCTTAAACCGTCTCGCTGAAGTCGGTAAGATCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAAACGGGATAAAGGCTTAAACCGTCTCGCTGAAGTCGGTAAGATCGGTAAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCGGATAAATCTGGAAGAGAAATTCGCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCGGATAAATCTGGAAGAGAAATTCGCCACAGGTT 1680
QY 1681 GCGGCAATCGGCGATGGCCCTGACATTAATCTTGGGCACACGCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAATCGGCGATGGCCCTGACATTAATCTTGGGCACACGCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTTGGCTGAAATCAACCCGGACAAAGCGTTCCAGGACAAAGCTGAT 1800
Db 1741 GCTCAATCTGGCTGTTGGCTGAAATCAACCCGGACAAAGCGTTCCAGGACAAAGCTGAT 1800
QY 1801 CCGTTTACCTGGATGCGGTACGTTACAAACCGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGGTACGTTACAAACCGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAAACCGCCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAAACCGCCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGGAAAGTAAGAGCGGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGGAAAGTAAGAGCGGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCTGATCTTCAACCTGCGCTGATTGCTGTCGACGGGGTATTGCGTTCAAG 2040
Db 1981 CTGCAAGAACCTGATCTTCAACCTGCGCTGATTGCTGTCGACGGGGTATTGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGATTAAGAACCTGGCGCTGGATTAACGCTGGCGGCAAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGATTAAGAACCTGGCGCTGGATTAACGCTGGCGGCAAAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACAACATGATGATGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACAACATGATGATGAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTAATAAAGCGAAACAGCGATGACCATCAACGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAAGCGAAACAGCGATGACCATCAACGCGCCGCTGG 2220
QY 2221 GCATGGTCCAAACATCGACACAGCAAAAGTGAATTTATGGTGAACGGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCAAACATCGACACAGCAAAAGTGAATTTATGGTGAACGGTACTGCCGACCTTC 2280
QY 2281 AAGGGTCAACCATCAACCGTTGCTGGTGGCTGCTGAGCGCAGGTATTACGCGCCGAGT 2340
Db 2281 AAGGGTCAACCATCAACCGTTGCTGGTGGCTGCTGAGCGCAGGTATTAAACGCGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTCTCTCGAAACATCTCTGCTGATGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTCTCTCGAAACATCTCTGCTGATGATGAAGTCTG 2400
QY 2401 GAACGGTTAATAAAGCAAAACCGCTGGTGGCTAGCTGAGTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAACGGTTAATAAAGCAAAACCGCTGGTGGCTAGCTGAGTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGCGAAAGATCCACGTTATGCGCCACCAATGGAACACCGCCCAAAAGGTGAATCATG 2520
Db 2461 TTGCGAAAGATCCACGTTATGCGCCACCAATGGAACACCGCCCAAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCGCGAGATGTCGCTTTCTGGTATGCCGCTGCTGCTGCGGTGATCAACGCC 2580
```

```
Db 2521 CCGAACATCCGCGAGATGTCGCTTTCTGGTATGTCGCTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGGTGCTCAGACTGTCGATGAAGCCCTGAAACAGCGCGCAGACTAATTCGAGTCTG 2640
Db 2581 GCCAGCGGTGCTCAGACTGTCGATGAAGCCCTGAAACAGCGCGCAGACTAATTCGAGTCTG 2640
QY 2641 AACAAACAACAATAAACAATAACAACAACTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACAAACAACAATAAACAATAACAACAACTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
QY 2701 GGATCTCTTCTCTGTGGCCAGCGCCCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTGGCCAGCGCCCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTTGATGTCCTTCAGCCAGCAGCTCCCTGCTGGTGGCCACACAGCGTATCCAC 2820
Db 2761 CCGGAATGTTGATGTCCTTCCTCTCAGAGCTCTCACCTGCTGGCCACACAGCGTATCCAC 2820
QY 2821 ACCGGTGAACAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACCGGTGAACAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGGTGGCCCATCAACGCACTCATATATGCGGAGAGCCATACAAATGTCAGAAATGTGGC 2940
Db 2881 CTGGTGGCCCATCAACGCACTCATATATGCGGAGAGCCATACAAATGTCAGAAATGTGGC 2940
QY 2941 AAGTCTTTTACGCGAGTCCAGCAACCTGGTGGCGCCACCAAGTACTCACCGGGGAGAAG 3000
Db 2941 AAGTCTTTTCTCGTCTGCAAAATCTCGTCCGGCACCAAGTACTCACCGGGGAGAAG 3000
QY 3001 CCTATGCTTTGTCGGGAATGTTGTAAGTCTTTCAGCACCAAGTGGCTCCTTTGGTTAGACAC 3060
Db 3001 CCTATGCTTTGTCGGGAATGTTGTAAGTCTTTCAGCGCAGCGATTAACCTGGTGGCCAC 3060
QY 3061 CAGGTATCCACACAGCGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
Db 3061 CAGGTATCCACACAGCGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
QY 3121 CAGCGGCCACCTGGAAACCCATCAACGCACTCATATGCGGAGAGCCATACAAATGT 3180
Db 3121 CAGCGGCCACCTGGCGAGCCATCAACGCACTCATATGCGGAGAGCCATACAAATGT 3180
QY 3181 CCAGAAATGTGGCAAGTCTTTCTCAACTTCAGGCAAACTTGGTCCGTCAACACCTACTCAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTTCTCTGCTCTGCAAACTCTGTCGCGCACCAACCTACTCAC 3240
QY 3241 ACCGTTAAAAAATACTAGTGGCCAGCGCCAGTACCGGTACGAGCTTCGGACTAGCT 3300
Db 3241 ACCGTTAAAAAATACTAGTGGCCAGCGCCAGTACCGGTACGAGCTTCGGACTAGCT 3300
```

RESULT 4

```
AX202427 LOCUS AX202427 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 15 from Patent WO0152620.
ACCESSION AX202427
VERSION AX202427.1 GI:15392175
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 15 26-JUL-2001;
The Scripps Research Institute (US); SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES
    source
        1. .3300
            /organism="synthetic construct"
            /mol_type="unassigned DNA"
            /db_xref="taxon:32630"
            /note="Partial sequence of pMal-m2 and zinc finger protein
```

ORIGIN	ZFPm2"	Db	961
Query Match Best Local Similarity 97.8%; Score 3226.4; DB 6; Length 3300; Matches 3254; Conservative 0; Mismatches 46; Indels 0; Gaps 0;			GCCTGCTGGGGCAAAACCGAGCTGGGACCCGCTTGCTGTCGAACCTCTCTCAGGGCCAGGCGGTGA 1020
		Qy	AGGGCAATCAGCTGTTGGCCCGCTCTCAGCTGGTGAAGAAAAAACCCACCTCTGGGGCCCAATA 1080
		Db	AGGGCAATCAGCTGTTGGCCCGCTCTCAGCTGGTGAAGAAAAAACCCACCTCTGGGGCCCAATA 1080
		Qy	CGCAAAACCGCTCTCCCGCGGTTGGCCCGATTCATTATGAGCTGGGACGACAGGTTT 1140
		Db	CGCAAAACCGCTCTCCCGCGGTTGGCCCGATTCATTATGAGCTGGGACGACAGGTTT 1140
		Qy	CCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTAAG 1200
		Db	CCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTAAG 1200
		Qy	GCACAAATTCATGTTTGCAGCTTATCATCGACTGCACGGTGACCAATGCTTCTGGCG 1260
		Db	GCACAAATTCATGTTTGCAGCTTATCATCGACTGCACGGTGACCAATGCTTCTGGCG 1260
		Qy	TCAGGAGCCATCGGAAGCTGTGGTATGCGTGTGAGTTCGTCGACCGTGCACCAATGCTTCTGGCG 1320
		Db	TCAGGAGCCATCGGAAGCTGTGGTATGCGTGTGAGTTCGTCGACCGTGCACCAATGCTTCTGGCG 1320
		Qy	TGTCGCTCAAGCGCACTCCCGTTCGGAATAATGTTTTTTCGCGCGACATCATACCGGTT 1380
		Db	TGTCGCTCAAGCGCACTCCCGTTCGGAATAATGTTTTTTCGCGCGACATCATACCGGTT 1380
		Qy	CTGGCAAAATATCTGAATGAGCTGTGTGACAAATTAATCATCGGCTCGTAAATGCTGTGA 1440
		Db	CTGGCAAAATATCTGAATGAGCTGTGTGACAAATTAATCATCGGCTCGTAAATGCTGTGA 1440
		Qy	ATTGTGAGCGGATPAAACAATTTTCACAGGAAACAGCCAGCTCCGTTTATGAGTGTGTTTTCACGA 1500
		Db	ATTGTGAGCGGATPAAACAATTTTCACAGGAAACAGCCAGCTCCGTTTATGAGTGTGTTTTCACGA 1500
		Qy	GCATTTCCACCAACAGGACCATAGATTATGAANAATGAAGAGGTAACCTGGTAAATCTCG 1560
		Db	GCATTTCCACCAACAGGACCATAGATTATGAANAATGAAGAGGTAACCTGGTAAATCTCG 1560
		Qy	ATTAAACCGGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAAATTCGAGAAAGAT 1620
		Db	ATTAAACCGGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAAATTCGAGAAAGAT 1620
		Qy	ACCGGAATTAAGTCAACGTTGAGCATCCGGATPAACTCGGAAGAGAAAATTCGCCACAGGTT 1680
		Db	ACCGGAATTAAGTCAACGTTGAGCATCCGGATPAACTCGGAAGAGAAAATTCGCCACAGGTT 1680
		Qy	CGGCAACTCGGCGATGGCCCTGCATTTCTCTGGGACACGACCGCTTGGTGCTAC 1740
		Db	CGGCAACTCGGCGATGGCCCTGCATTTCTCTGGGACACGACCGCTTGGTGCTAC 1740
		Qy	GCTCAATCTGGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
		Db	GCTCAATCTGGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
		Qy	CCGTTTACCTGGGATCCGTTACACCGCAAGCTGATTGCTTACCAGTCCGCTGTT 1860
		Db	CCGTTTACCTGGGATCCGTTACACCGCAAGCTGATTGCTTACCAGTCCGCTGTT 1860
		Qy	GAAGCGTTATCGCTGATTTTATAAAGATCTGCTCCGAAACCCGCCAAAAAACCTGGGAA 1920
		Db	GAAGCGTTATCGCTGATTTTATAAAGATCTGCTCCGAAACCCGCCAAAAAACCTGGGAA 1920
		Qy	GAGATCCCGCGCTGATATAAGAACTGAAAGCGGAAAGGTAAGAGCGCGCTGATGTTTCAAC 1980
		Db	GAGATCCCGCGCTGATATAAGAACTGAAAGCGGAAAGGTAAGAGCGCGCTGATGTTTCAAC 1980
		Qy	CTGCAAGAACCGTACTTACCTGGCCGCTGATTGCTGTGACGGGGGTTATGCGTTTCAAG 2040
		Db	CTGCAAGAACCGTACTTACCTGGCCGCTGATTGCTGTGACGGGGGTTATGCGTTTCAAG 2040
		Qy	TATGAAAAACCGCAAGTACGACATTAAGAAGCTGGGCGTGGATTAACGCTGGCGGAAAGCG 2100
		Db	TATGAAAAACCGCAAGTACGACATTAAGAAGCTGGGCGTGGATTAACGCTGGCGGAAAGCG 2100

```
QY 2101 GGTCTGAACCTTCTCTGTTGAACCTGATTAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGAACCTTCTCTGTTGAACCTGATTAAACAAACACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTAAATAAAGCGGAAACAGCGATGACCATCAACGGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTAAATAAAGCGGAAACAGCGATGACCATCAACGGCCGCTGG 2220
QY 2221 GCATGGTCAACATCGACACACAGCAAAAGTGAATTAATGGTTAAACGGTACTCCGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACACAGCAAAAGTGAATTAATGGTTAAACGGTACTCCGACCTTC 2280
QY 2281 AAGGTTCAACCATCAAAACCGTTGCTGGCGGTGTCAGCGCAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTTCAACCATCAAAACCGTTGCTGGCGGTGTCAGCGCAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGGAAAGATCCAACGTTATTTGGCGCCACCATGGAACACGCCCAAGAGGTGAATCATG 2520
Db 2461 TTGGGAAAGATCCAACGTTATTTGGCGCCACCATGGAACACGCCCAAGAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
QY 2581 GCCAGCGTCTGAGCTGTGCTGAGAGCCCTGAAAGACGCGCAGACTAATTCAGGCTCG 2640
Db 2581 GCCAGCGTCTGAGCTGTGCTGAGAGCCCTGAAAGACGCGCAGACTAATTCAGGCTCG 2640
QY 2641 AACAAACAAACAAATACAAATACAAACAACTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAAACAAATACAAATACAAACAACTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
QY 2701 GGATCTCTTCTCTGTCGCGCCAGCGCCCTCCAGCCCGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTCGCGCCAGCGCCCTCCAGCCCGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGAAATGTGTAACTCTTCAGCAGCAGAGCTCCCTGCTGCGCCACACGCGTACCCAC 2820
Db 2761 CCGAAATGTGTAACTCTTCAGAGCTCTCACTGCTGCGCCACACGCGTACCCAC 2820
QY 2821 ACGGTGAACACCGTATAATGCGCAGAGTGGCGGAAATCTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGTGAACACCGTATAATGCGCAGAGTGGCGGAAATCTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGGTGGCCCATCAACGCACTCATCTGCGGAGAGGCATACAATGTCCAGAAATGTCG 2940
Db 2881 CTGGTGGCCCATCAACGCACTCATCTGCGGAGAGGCATACAATGTCCAGAAATGTCG 2940
QY 2941 AAGTCTTTCTCGGTCTGACAACTCTGTCGCGCACCAAGTACTCAACCGGGGAGAAG 3000
Db 2941 AAGTCTTTCTCGGTCTGACAACTCTGTCGCGCACCAAGTACTCAACCGGGGAGAAG 3000
QY 3001 CCCTATGCTTGTTCGGAAATGTTGTAAGTCTTTCAGACCAAGTGTCTTGGTTAGACAC 3060
Db 3001 CCCTATGCTTGTTCGGAAATGTTGTAAGTCTTTCAGACCAAGTGTCTTGGTTAGACAC 3060
QY 3061 CAGCGTACCACACCGGTGAAACCGTATAAATGCGCAGAGTGGCGCAAACTCTTTTACG 3120
Db 3061 CAGCGTACCACACCGGTGAAACCGTATAAATGCGCAGAGTGGCGCAAACTCTTTTACG 3120
QY 3121 CAGCGGCCCACTGTGAACCGCATCAACGCACTCATCTGCGGAGAGCGCATCAAAATGT 3180
Db 3121 CAGCGGCCCACTGTGGCCAGCCATCAACGCACTCATCTGCGGAGAGCGCATCAAAATGT 3180
```

```
QY 3181 CCAGAAATGTGGCAAGTCTTTTCTCAACTTCAGGCAACTTTGGTCGCTACCAACGTAACAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTTTCTCGGCTTGACAAATCTCGTCCGCAACCAACGTAACAC 3240
QY 3241 ACCGGTAAAAAATACTAGTGGCCAGCGCGGCGGACGTAACCGTACGACGTTCCGGACTACGCT 3300
Db 3241 ACCGGTAAAAAATACTAGTGGCCAGCGCGGCGGACGTAACCGTACGACGTTCCGGACTACGCT 3300

RESULT 5
AX202428 LOCUS 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 16 from Patent WO0152620.
ACCESSION AX202428
VERSION AX202428.1 GI:15392176
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas, C.F., Stege, J.T., Guan, X. and Dalmia, B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 16 JUL-2001;
The Scripps Research Institute (US); SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES
source 1..3300
location/qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="Partial sequence of pMal-m3 and zinc finger protein ZFPm3"
```

ORIGIN

```
Query Match 97.5%; Score 3216.8; DB 6; Length 3300;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 3248; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGTCGCAAAACCTTTCCGCGTATGGCATGATAGCGCCCGGAGAGA 60
Db 1 CCGACACCATCGAATGTCGCAAAACCTTTCCGCGTATGGCATGATAGCGCCCGGAGAGA 60
QY 61 GTCATTCAGGTTGGTGAATGTGAACACCAAGTAACTATACGATGTCGAGAGTATGCGG 120
Db 61 GTCATTCAGGTTGGTGAATGTGAACACCAAGTAACTATACGATGTCGAGAGTATGCGG 120
QY 121 GTCTCTTTATCAGACCGTTTCCCGCGTGGTGAACAGCGCCAGCCAGCTTCTTCGCAAAA 180
Db 121 GTCTCTTTATCAGACCGTTTCCCGCGTGGTGAACAGCGCCAGCCAGCTTCTTCGCAAAA 180
QY 181 CCGGGGAAAAAGTGGAAAGCGGCGATGGCGGAGCTGAATTAACATTCACACCGCGTGGCAC 240
Db 181 CCGGGGAAAAAGTGGAAAGCGGCGATGGCGGAGCTGAATTAACATTCACACCGCGTGGCAC 240
QY 241 AACAACTGGCGGCAACAGTCTGTTCTGATTGGCGTTGCCACTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAACAGTCTGTTCTGATTGGCGTTGCCACTCCAGTCTGGCCCTGC 300
QY 301 ACGCCGCTGGCAAAATGTCGGCGCATTAATATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCCGCTGGCAAAATGTCGGCGCATTAATATCTCGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGTGTGATGTTAGAACGAGCGCGTGAAGCCCTGTAAAGCGGCGGTGCACAATC 420
Db 361 TGGTGTGTGATGTTAGAACGAGCGCGTGAAGCCCTGTAAAGCGGCGGTGCACAATC 420
QY 421 TTCTCCGCAACCGCTCAGTGGGTGATCATTAATTAATCTCCGCTGGATGACCAAGATGCCA 480
Db 421 TTCTCCGCAACCGCTCAGTGGGTGATCATTAATTAATCTCCGCTGGATGACCAAGATGCCA 480
QY 481 TTGCTGTGAAGCTGCTGCACCTAAATGTTCCGCGCTTATTTCTTGATGTCTCTGACGAGA 540
Db 481 TTGCTGTGAAGCTGCTGCACCTAAATGTTCCGCGCTTATTTCTTGATGTCTCTGACGAGA 540
```


QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC 600
QY 601 TGGTCCGATTCGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db 601 TGGTCCGATTCGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGCTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCGATAG 720
Db 661 CGCGTCTCGCTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCGATAG 720
QY 721 CGGAACGGGAAGGACCTGGAGTGCATGTCGGTTTTCAACAACCAATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGACCTGGAGTGCATGTCGGTTTTCAACAACCAATGCAAAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTCGGATGCTGGTTGCCAAGATCAGATGCGCTGGGCGCAA 840
Db 781 ATGAGGGCATGTTCCCACTCGGATGCTGGTTGCCAAGATCAGATGCGCTGGGCGCAA 840
QY 841 TCGCGCCATTACCGAGTCCGGCTCGCGCTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TCGCGCCATTACCGAGTCCGGCTCGCGCTTGGTGGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCGCCGTTAAACCAACATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCGCCGTTAAACCAACATCAAAACAGGATTTTC 960
QY 961 GCCTGTGGGGCAACACAGCTGGAACCGCTTGTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGTGGGGCAACACAGCTGGAACCGCTTGTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGCACTGGCGACGACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAAATGCACTGGCGACGACAGGTTT 1140
QY 1141 CCGGACTGGAAGCGGCAGTGAGCGCAACCAATTAATGTAGTTAGTCACTCATTTAG 1200
Db 1141 CCGGACTGGAAGCGGCAGTGAGCGCAACCAATTAATGTAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGTGAACCAATTAATGTAGTTAGTCACTCATTTAG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGTGAACCAATTAATGTAGTTAGTCACTCATTTAG 1260
QY 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTGAATCACTGCAATTAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTGAATCACTGCAATTAATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCGCTTCTGGATATGTTTTTGGCGGACATCATACCGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCGCTTCTGGATATGTTTTTGGCGGACATCATACCGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
QY 1441 ATTGTAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTTAGGTTTTTCAAGA 1500
Db 1441 ATTGTAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTTAGGTTTTTCAAGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAACCTGAAGAGGTAAACCTGATCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTATGAACCTGAAGAGGTAAACCTGATCTGG 1560
QY 1561 ATTTAAGCGGATAAAGGCTTAAACGCTCTGCTGAAGTCCGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTTAAGCGGATAAAGGCTTAAACGCTCTGCTGAAGTCCGTAAGAAATTCAGAAAGAT 1620

QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCCGATAAACTGGAAGAGAAATTTCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCCGATAAACTGGAAGAGAAATTTCCACAGGTT 1680
QY 1681 GCGCAAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTGGCTGAAATCAACCCCGAACAAGGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTGGCTGAAATCAACCCCGAACAAGGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGTACGTTACAAACCGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGTACGTTACAAACCGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
QY 1861 GAACGCTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAACGCTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAAGCGAAAGTAAAGAGCGCGCTGATTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAAGCGAAAGTAAAGAGCGCGCTGATTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGAACCGTGGCGCTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGAACCGTGGCGCTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTGGTTGACCTGATTAATAAAACAAACAATGAATGCAACCGGATTTAC 2160
Db 2101 GGTCTGACCTTCTCGTGGTTGACCTGATTAATAAAACAAACAATGAATGCAACCGGATTTAC 2160
QY 2161 TCCATCCGAGAGCTGCTTTTAATAAGGCGAAACAGCGATGACCATCAAGCGGCCGCTGG 2220
Db 2161 TCCATCCGAGAGCTGCTTTTAATAAGGCGAAACAGCGATGACCATCAAGCGGCCGCTGG 2220
QY 2221 GCATGCTCAACATCGACACCAACAGTGAATTAATGTTGTAACCGTACTGCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACCAACAGTGAATTAATGTTGTAACCGTACTGCGGACCTTC 2280
QY 2281 AAGGCTCAACCATCAAAACCGTTCTGTTGGCGTGTGAGCGCAGGTTAAACCGCCAGT 2340
Db 2281 AAGGCTCAACCATCAAAACCGTTCTGTTGGCGTGTGAGCGCAGGTTAAACCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGCCTGAGCGCTGAACTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGCCTGAGCGCTGAACTCTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCACGTTATTGCGGCCCACTTGAAGAAAACGCCCGAAGAGGTAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATTGCGGCCCACTTGAAGAAAACGCCCGAAGAGGTAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGTCGCTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGTCGCTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACCAACAAACAAATAACAAACCAACCTCGGGATCGAGGGAAGGATTTTCAAGATTC 2700
Db 2641 AACCAACAAACAAATAACAAACCAACCTCGGGATCGAGGGAAGGATTTTCAAGATTC 2700
QY 2701 GGATCCTCTTCTGTTGGCCAGCGCGGCCCTCGAGCGCGGGGAGAGACCCCTATGCTGT 2760


```
Db 2701 GGATCCTCTTCTCTGTGGCCAGCGCGCCCTCGAGCCGCGGGAAGACCCCTATGCTTGT 2760
Qy 2761 CCGGAATGTGGTAAGTCTCTAGCCAGAGCAGCTCCCTGGTGGCCACACAGGTTACCCAC 2820
Db 2761 CCGGAATGTGGTAAGTCTCTAGCCAGATCTGGTGGCCACCTGGTGGCCACACAGGTTACCCAC 2820
Qy 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAACTTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAACTTTTAGCCAGGCGCTCC 2880
Qy 2881 CTGGTGGCCATCAACGCACTCTACTCTGGGAGAGCCATCAAAATGTCAGAGTGTGGC 2940
Db 2881 CTGGTGGCCATCAACGCACTCTACTCTGGGAGAGCCATCAAAATGTCAGAGTGTGGC 2940
Qy 2941 AAGTCTTTTTCAGCAGTCCAGCAACCTGGTGGCCACCAACCTACTCTACACCGGGAGAG 3000
Db 2941 AAGTCTTTTTCAGCAGTCCAGCCTGGTGGCCACCAACCTACTCTACACCGGGAGAG 3000
Qy 3001 CCTATGCTTGTCCGGAATGTGGTAAGTCTTTCAGCAGCAGTGGCTCCTTGGTTAGACAC 3060
Db 3001 CCTATGCTTGTCCGGAATGTGGTAAGTCTTTCAGCAGCAGCAGTCCCTCGTGGCCAC 3060
Qy 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAACTTTTAGC 3120
Db 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAACTTTTAGT 3120
Qy 3121 CAGCGCGCCACCTGGAAACGGCATCAACGCACTCTACTCTGGCGAGAGCCATCAAAATGT 3180
Db 3121 GACTGCCGCGACCTTGTCTGCCATCAACGCACTCTACTCTGGCGAGAGCCATCAAAATGT 3180
Qy 3181 CCAGATGTGGCAAGTCTTTCTCACTTCAGGCAACTTGGTCCGTCAACCACTACTCAC 3240
Db 3181 CCAGATGTGGCAAGTCTTTCTCCAAATCCAGCCATCTCGTCCGCGCAACCACTACTCAC 3240
Qy 3241 ACCGTAAGAAAACTAGTGGCGAGCGCGCCAGTACCGGTACGAGCTTCGGACTACGCT 3300
Db 3241 ACCGTAAGAAAACTAGTGGCGAGCGCGCCAGTACCGGTACGAGCTTCGGACTACGCT 3300

RESULT 6
AX284144 LOCUS
DEFINITION Sequence 1 from Patent WO0178756.
ACCESSION AX284144
VERSION AX284144.1 GI:17044832
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1.
AUTHORS Wiedersanders, B. and Maubach, G.
TITLE Agent for postoperative use after the removal of bone tumours
JOURNAL Patent: WO 0178756-A 1 25-OCT-2001;
Depuy Biotech Jena GmbH (DE)
FEATURES
source
location/Qualifiers
1..7373
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Expressionsvektor pMal2cchbnp2"
1528..3453
/notes="unnamed protein product; /produkt=Fusionsprotein
MBP Cystatin C BMP-2"
/codon_start=1
/transl_table=11
/protein_id="CAD12160.1"
/db_xref="GI:17044833"
/translation="MKIEBGLVIWINGDKYNGLAIEVPGKFKDTGIKVTVEHPDKL
EEKFPQVAATGDPDIIIFWADRFGRGVAQSGLLAEITPDRAFDQKLYPTFTWADVRNG
KLAIYAPFAEALSLIYNKOLLNPNPKTWEEIIPALDKELKAKGKLSALMFLNLOEYFTWP
LIAADGGYAPKENGKVDI KDVGVDNAGAKGLTFLVDLTKNKHMDATDYSIAEAP
NKGETAWTINGPNAWSNIDTSKYNIGVTVLPTFKGQPSKFPFVGVLNAGINAASPNKEL
```

```
misc_feature
/note="Faktor Xa Protease /position=388 (AMINO ACIDsequenz)"
3067..3108
/misc_feature
/note="/spacer zwischen Cystatin C und BMP-2"
ORIGIN
Query Match 82.0%; Score 2705.8; DB 6; Length 7373;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2719; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGCAAAACCTTTCGGGTATGGCATATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTCGGGTATGGCATATAGCGCCCGGAAGAGA 60
Qy 61 GTCAATTCAGGGTGGTGAATGTCAAAACCTTATACGATGTGCGAGAGTATCGCG 120
Db 61 GTCAATTCAGGGTGGTGAATGTCAAAACCTTATACGATGTGCGAGAGTATCGCG 120
Qy 121 GTGTCTCTTATCAGACCGTTTCCGGGTGTGTAACAGGCGCAGCAGTTCCTGCGAAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCGGGTGTGTAACAGGCGCAGCAGTTCCTGCGAAAA 180
Qy 181 CCGGGGAAAAAAGTGGAAAGCGCGCATGGCGGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 240
Db 181 CCGGGGAAAAAAGTGGAAAGCGCGCATGGCGGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 240
Qy 241 AACAACTGGCGGGGAAACAGTCTGCTGATTTGGCGTTCGCACTCAAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGGAAACAGTCTGCTGATTTGGCGTTCGCACTCAAGTCTGGCCCTGC 300
Qy 301 ACCGGCGCTCGCAAAATGTGGCGGCAATTAATCTCGCGCGGATCAACTGGGTGGCAGG 360
Db 301 ACCGGCGCTCGCAAAATGTGGCGGCAATTAATCTCGCGCGGATCAACTGGGTGGCAGG 360
Qy 361 TGGTGGTGTGATGGTAGAACGAGCGGCTGCAAGCCTGTAAGCGCGCGGTGCACAATC 420
Db 361 TGGTGGTGTGATGGTAGAACGAGCGGCTGCAAGCCTGTAAGCGCGCGGTGCACAATC 420
Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACAGGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCTGCACATAATTTTCGGCGGTATTTCTTGTATGCTCTGACGAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACATAATTTTCGGCGGTATTTCTTGTATGCTCTGACGAGA 540
Qy 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Qy 601 TGGTGGCATTTGGGTACACAGCAAAATCGCGTGTAGCGGGCCCATTAAGTTCCTGCTCGG 660
Db 601 TGGTGGCATTTGGGTACACAGCAAAATCGCGTGTAGCGGGCCCATTAAGTTCCTGCTCGG 660
Qy 661 CCGGTCTGCGTCTGGGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCCATAG 720
Db 661 CCGGTCTGCGTCTGGGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCCATAG 720
Qy 721 CGGAACGGGAAGCGACTGGAGTGGCATGTCCGGTTTTTCAACAAACCAATGCAAAATCTGA 780
Db 721 CGGAACGGGAAGCGACTGGAGTGGCATGTCCGGTTTTTCAACAAACCAATGCAAAATCTGA 780
Qy 781 ATGAGGGCATCTGTTCCCATCGGATGCTGTTGCCACCATCAGATGGGCGCTGGGGCGAA 840
Db 781 ATGAGGGCATCTGTTCCCATCGGATGCTGTTGCCACCATCAGATGGGCGCTGGGGCGAA 840
```

QY 841 TCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TCGCGCCATTACCGAGTCCGGGCTGCGCGTTGGTGGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCAACATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCAACATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGGCAAAACCGAGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGGCAAAACCGAGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCGCTCTCATCTGGTGAAGAAAAACCAACCTCTGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCGCTCTCATCTGGTGAAGAAAAACCAACCTCTGCGGCCAATA 1080
QY 1081 CGCAACCGCCCTCTCCCGCGGTTGGCCGATTCATTAATGCGTGGCAACAGGTTT 1140
Db 1081 CGCAACCGCCCTCTCCCGCGGTTGGCCGATTCATTAATGCGTGGCAACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGGCGCAACGCAATTAATGTAGTCTCACTCACTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGGCGCAACGCAATTAATGTAGTCTCACTCACTTAG 1200
QY 1201 GCACAATTTCTCATGTTGACAGCTTATCATGCTGCAAGGCAACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTTCTCATGTTGACAGCTTATCATGCTGCAAGGCAACCAATGCTTCTGGCG 1260
QY 1261 TCAGGCGCCATCGAAGCTGGTATGGCTGCGAGTGGTAAATCACTGCAATTAATTCG 1320
Db 1261 TCAGGCGCCATCGAAGCTGGTATGGCTGCGAGTGGTAAATCACTGCAATTAATTCG 1320
QY 1321 TGTGCTCAAGCGCATCTCCGCTTCTGGATAATGTTTTTGGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGCATCTCCGCTTCTGGATAATGTTTTTGGCGCGACATCAATACGGTT 1380
QY 1381 CTGGCAATTTCTGAATGAGCTGTTGACATTAATCAATCAATCGGCTCGTAAATTCG 1440
Db 1381 CTGGCAATTTCTGAATGAGCTGTTGACATTAATCAATCAATCGGCTCGTAAATTCG 1440
QY 1441 ATTGTAGCGGATTAACATTTACACAGGAACAGCCAGTCCGTTAGTGTTCACGA 1500
Db 1441 ATTGTAGCGGATTAACATTTACACAGGAACAGCCAGTCCGTTAGTGTTCACGA 1500
QY 1501 GCACTTCCACCAAGGACCATAGATTATGAAAACTGAAGAAGTTAACTGGTAAATCTGG 1560
Db 1501 GCACTTCCACCAAGGACCATAGATTATGAAAACTGAAGAAGTTAACTGGTAAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACCGTCTCGTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAAAGCTCTCGTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTAGGCTGAGTAACTCGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTAGGCTGAGTAACTCGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTGGGCAACAGCCGCTTCTGGTGTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTGGGCAACAGCCGCTTCTGGTGTAC 1740
QY 1741 GCTCAATCTGGCTGTTGGCTGAATCAACCCCGGCAACAGGTTCCAGGCAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTTGGCTGAATCAACCCCGGCAACAGGTTCCAGGCAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCGGTAGCTTAACTCAACCGGCAACAGGTTCCAGGCAAGCTGTAT 1860
Db 1801 CCGTTTACCTGGATGCGGTAGCTTAACTCAACCGGCAACAGGTTCCAGGCAAGCTGTAT 1860
QY 1861 GAAGCGTTATCGCTGATTTTATACAAAGATCTGCTGCCAAACCCGCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTTATACAAAGATCTGCTGCCAAACCCGCAAAACCTGGGAA 1920

QY 1921 GAGATCCGGCGCTGGATAAAGAACTGAAGCGAAAGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGGCGCTGGATAAAGAACTGAAGCGAAAGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAAACGTAATCTTCACTGCGCGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAAACGTAATCTTCACTGCGCGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGAACCTGCGCGCTGGATAAACCTGCGCGCAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGAACCTGCGCGCTGGATAAACCTGCGCGCAAGCG 2100
QY 2101 GGTCTGACCTTCTGCTGATTAATAAACAACAACAACAATGATGACAGACCGATTAAC 2160
Db 2101 GGTCTGACCTTCTGCTGATTAATAAACAACAACAACAATGATGACAGACCGATTAAC 2160
QY 2161 TCCATCGCAGAAGCTGCTTTTAAAGCGCAACAGCATGACCATCAAGCGGCGCGTGG 2220
Db 2161 TCCATCGCAGAAGCTGCTTTTAAAGCGCAACAGCATGACCATCAAGCGGCGCGTGG 2220
QY 2221 GCATGGTCCAAATCGACACCAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACCAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGTCAACCATCCAAACCGTTGTTGGCGTCTGAGCGAGGTAATTAACCGCGCCAGT 2340
Db 2281 AAGGTCAACCATCCAAACCGTTGTTGGCGTCTGAGCGAGGTAATTAACCGCGCCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAACATATCTGCTGATGATGAAGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAACATATCTGCTGATGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAAGACAAACCGCTGGTCCGCTAGCGCTGAAGTCTTACGAGAAAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCGCTGGTCCGCTAGCGCTGAAGTCTTACGAGAAAG 2460
QY 2461 TTGGCAAAAGATCCAGTATTTGGCGCCACCATGGAACCGCCAGAAAGGTAATCATG 2520
Db 2461 TTGGCAAAAGATCCAGTATTTGGCGCCACCATGGAACCGCCAGAAAGGTAATCATG 2520
QY 2521 CCGAACATCCCGAGATGCTCGCTTTCTGCTGATGCTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACATCCCGAGATGCTCGCTTTCTGCTGATGCTGCTGCTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTG 2640
Db 2581 GCCAGCGTCTGACGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTG 2640
QY 2641 AACAAACAACAATAAACAATAACAACACCTCGGGATCGAGGAAAGGATTTCAAGATTC 2700
Db 2641 AACAAACAACAATAAACAATAACAACACCTCGGGATCGAGGAAAGGATTTCAAGATTC 2700
QY 2701 GGATCCTCTCTCTGCTGGCGCCAGCGCCCTTCGAGCGCCG 2741
Db 2701 GGATCCTCTCTCTGCTGGCGCCAGCGCCCTTCGAGCGCCG 2741

RESULT 7
AX172306
LOCUS AX172306 8101 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 5 from Patent WO0144467.
ACCESSION AX172306
VERSION AX172306.1 GI:14597487
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Goubin-Gramatica,F., Ducommun,B. and Prevost,G.
TITLE Method for obtaining human cdc25 phosphatases and method for
identifying human cdc25 phosphatase modulators
JOURNAL Patent: WO 0144467-A 5 21-JUN-2001;
SOCIETE DE CONSEILS DE RECHERCHES ET D'APPLICATIONS SCIENTIFIQUES

FEATURES		(PR)		Location/Qualifiers	
source		1..8101		/organism="synthetic construct"	
				/mol_type="unassigned DNA"	
				/db_xref="taxon:32630"	
				/note="souche JM109 / pMAL - Hs Cdc25C"	
ORIGIN					
Query Match		82.0%;		Score 2705.8; DB 6; Length 8101;	
Best Local Similarity		99.9%;		Pred. No. 0;	
Matches 2707;		Conservative 0;		Mismatches 2; Indels 0; Gaps 0;	
Qy	1	CCGACACCATCGAATGGTGTGCAAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA	60		
Db	1	CCGACACCATCGAATGGTGTGCAAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA	60		
Qy	61	GTCAATTCAGGGTGGTGAATGTGAACCAAGTAAAGTTATACGATGTGCGAGTATGCGG	120		
Db	61	GTCAATTCAGGGTGGTGAATGTGAACCAAGTAAAGTTATACGATGTGCGAGTATGCGG	120		
Qy	121	GTGCTCTTATCAGACCGTTTCCCGCGTGTGAACCGAGCCAGCCAGCTTTCTGCGAAA	180		
Db	121	GTGCTCTTATCAGACCGTTTCCCGCGTGTGAACCGAGCCAGCCAGCTTTCTGCGAAA	180		
Qy	181	CGCGGAAAAAGTGGAGCGCGCATGGCGAGCTGAATTCATTCCCAACCGCGTGGCAC	240		
Db	181	CGCGGAAAAAGTGGAGCGCGCATGGCGAGCTGAATTCATTCCCAACCGCGTGGCAC	240		
Qy	241	AACAACTGGCGGCAACAGTCGTGTGATGTGGGTGGCCACCTCGATGTGGCCCTGC	300		
Db	241	AACAACTGGCGGCAACAGTCGTGTGATGTGGGTGGCCACCTCGATGTGGCCCTGC	300		
Qy	301	ACGCGCGCTCGCAATTTGTCCGCGCATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360		
Db	301	ACGCGCGCTCGCAATTTGTCCGCGCATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360		
Qy	361	TGGTGGTTCGATGGTGAAGCGCGCTGGAAGCGCTGTAAGCGCGCGTGCACAATC	420		
Db	361	TGGTGGTTCGATGGTGAAGCGCGCTGGAAGCGCTGTAAGCGCGCGTGCACAATC	420		
Qy	421	TTCTCGGCAACGGTTCAGTGGGTGATCATTAATCTCGCTGGATGACCGAGTGC	480		
Db	421	TTCTCGGCAACGGTTCAGTGGGTGATCATTAATCTCGCTGGATGACCGAGTGC	480		
Qy	481	TTGCTGTGGAAGCTGCCTGCAATAATGTTCCGGCGTTATTTCTTGATGTCCTCGACCAGA	540		
Db	481	TTGCTGTGGAAGCTGCCTGCAATAATGTTCCGGCGTTATTTCTTGATGTCCTCGACCAGA	540		
Qy	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600		
Db	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600		
Qy	601	TGGTGGCATGGGTCCACAGCAATCGCGTGTAGCGGGGCCATTAAAGTTCGTCCTCGG	660		
Db	601	TGGTGGCATGGGTCCACAGCAATCGCGTGTAGCGGGGCCATTAAAGTTCGTCCTCGG	660		
Qy	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720		
Db	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720		
Qy	721	CGGAACGGGAAGGCACTGGAGTCCCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA	780		
Db	721	CGGAACGGGAAGGCACTGGAGTCCCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA	780		
Qy	781	ATGAGGGCATCGTTCCCACTCGCATGCTGGTGTGCCAACAGATCAGATGCGCTGGCGCAA	840		
Db	781	ATGAGGGCATCGTTCCCACTCGCATGCTGGTGTGCCAACAGATCAGATGCGCTGGCGCAA	840		
Qy	841	TGCGCGCATTAACGAGTCCGGCTGCGGTGTGGATATCTCGGTAGTGGGATACG	900		
Db	841	TGCGCGCATTAACGAGTCCGGCTGCGGTGTGGGATATCTCGGTAGTGGGATACG	900		

Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGTTTAAACACCATCAAAACAGGATTTTC	960		
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGTTTAAACACCATCAAAACAGGATTTTC	960		
Qy	961	GCTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020		
Db	961	GCTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020		
Qy	1021	AGGCAATCAGCTGTGCGCGCTCTCACTGCTGGAAGAAAAACACCTGGCGCCCAATA	1080		
Db	1021	AGGCAATCAGCTGTGCGCGCTCTCACTGCTGGAAGAAAAACACCTGGCGCCCAATA	1080		
Qy	1081	CGCAAAACCGCTCTCTCCCGCGGTGGCCGATTCATTAAATGCACTGCGACGACGTTT	1140		
Db	1081	CGCAAAACCGCTCTCTCCCGCGGTGGCCGATTCATTAAATGCACTGCGACGACGTTT	1140		
Qy	1141	CCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTAG	1200		
Db	1141	CCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTAG	1200		
Qy	1201	GCACAAATTCATGTTTGACAGCTTATCATCGACTGCACTGCGTGCACCAATGCTTCTGGC	1260		
Db	1201	GCACAAATTCATGTTTGACAGCTTATCATCGACTGCACTGCGTGCACCAATGCTTCTGGC	1260		
Qy	1261	TCAGGCAAGCCATCGGAAGCTGTGTATGGCTGTGCAAGTGTAAATCACTGCAATAATTCG	1320		
Db	1261	TCAGGCAAGCCATCGGAAGCTGTGTATGGCTGTGCAAGTGTAAATCACTGCAATAATTCG	1320		
Qy	1321	TGTCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTTGGCGCGACATCAATAACGGT	1380		
Db	1321	TGTCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTTGGCGCGACATCAATAACGGT	1380		
Qy	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGCAAAATTAATCATCGCTCGTAAATGTTGGGA	1440		
Db	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGCAAAATTAATCATCGCTCGTAAATGTTGGGA	1440		
Qy	1441	ATTGTGAGCGGATAACAAATTTTCACAGGAAACAGCCAGTCCGTTTAGTGTGTTTTCACGA	1500		
Db	1441	ATTGTGAGCGGATAACAAATTTTCACAGGAAACAGCCAGTCCGTTTAGTGTGTTTTCACGA	1500		
Qy	1501	GCACTTCACCAACAGGACCATAGATTGAAATCTGAAGAAGTAAACTGGTAACTCTGG	1560		
Db	1501	GCACTTCACCAACAGGACCATAGATTGAAATCTGAAGAAGTAAACTGGTAACTCTGG	1560		
Qy	1561	ATTAAACGGGCATAAAGGCTATAACGGTCTCGCTGAAGTTCGTAAGAAATTCGAGAAAGAT	1620		
Db	1561	ATTAAACGGGCATAAAGGCTATAACGGTCTCGCTGAAGTTCGTAAGAAATTCGAGAAAGAT	1620		
Qy	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGTT	1680		
Db	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGTT	1680		
Qy	1681	GGCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACCGCTTTGGTGGCTAC	1740		
Db	1681	GGCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACCGCTTTGGTGGCTAC	1740		
Qy	1741	GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGCAAAAGCGTTCAGGACAAAGCTGTAT	1800		
Db	1741	GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGCAAAAGCGTTCAGGACAAAGCTGTAT	1800		
Qy	1801	CCGTTTACCTGGGATGCGGTACGTTTAAACAAAGATCTGCTGCCGAACCGCCAAAACCTGGAA	1860		
Db	1801	CCGTTTACCTGGGATGCGGTACGTTTAAACAAAGATCTGCTGCCGAACCGCCAAAACCTGGAA	1860		
Qy	1861	GAAGCGTTATCGCTGATTTTATCAAAAGATCTGCTGCCGAACCGCCAAAACCTGGAA	1920		
Db	1861	GAAGCGTTATCGCTGATTTTATCAAAAGATCTGCTGCCGAACCGCCAAAACCTGGAA	1920		
Qy	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGACGCGCTGTGTTTCAAC	1980		
Db	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGACGCGCTGTGTTTCAAC	1980		
Qy	1981	CTGCAAGAACCGTACTCTACCTGGCGCTGATTGCTGCTGACGGGGTTATGCTTCAAG	2040		

				/note="Plasmid pMAL-c2x"	
ORIGIN	Query Match Best Local Similarity Matches 2704; Conservative	81.8%; 99.8%; 0;	Score 2701; Pred. No. 0; Mismatches	DB 6; Length 6648; 5;	Indels 0; Gaps 0;
Qy	1	CCGACACCATCGAATGTCGAAAAACCTTTCCGCGGTATGGCATGATAGCCCGCGGAAGAGA	60		
Db	1	CCGACACCATCGAATGTCGAAAAACCTTTCCGCGGTATGGCATGATAGCCCGCGGAAGAGA	60		
Qy	61	GTCAATTACAGGTGCTCAATGTGAACCCAGTACGTTTATACGATGTGCGAGAGTATGCCG	120		
Db	61	GTCAATTACAGGTGCTCAATGTGAACCCAGTACGTTTATACGATGTGCGAGAGTATGCCG	120		
Qy	121	GTGTCCTCTTATCAGACCCGTTTCCCGCGTGTGAACCCAGGCCAGCCACCTGTTCTTCGCAAAA	180		
Db	121	GTGTCCTCTTATCAGACCCGTTTCCCGCGTGTGAACCCAGGCCAGCCACCTGTTCTTCGCAAAA	180		
Qy	181	CGCGGAAAAAGTGGAAAGCGCGAGTGGCGAGCTGAATTACATTCCCAACCCGCGTGGCAC	240		
Db	181	CGCGGAAAAAGTGGAAAGCGCGAGTGGCGAGCTGAATTACATTCCCAACCCGCGTGGCAC	240		
Qy	241	AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCAGTCTGGCCCTGC	300		
Db	241	AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCAGTCTGGCCCTGC	300		
Qy	301	ACGCGCCGTTCGCAAAATTTGTCGCGCGGATTAATATCTCGCGCCGATCAACTGGGTGCCAGCG	360		
Db	301	ACGCGCCGTTCGCAAAATTTGTCGCGCGGATTAATATCTCGCGCCGATCAACTGGGTGCCAGCG	360		
Qy	361	TGGTGGTGTTCGATGGTAGAAACGAAAGCGGCGTCAAGAGCGCTGTAAAGCGCGGTGCACAA	420		
Db	361	TGGTGGTGTTCGATGGTAGAAACGAAAGCGGCGTCAAGAGCGCTGTAAAGCGCGGTGCACAA	420		
Qy	421	TTCTCGCGCAACCGGTCAAGTGGGCTGATCAATTAATCTCGCTGGATGACCAAGATGCCA	480		
Db	421	TTCTCGCGCAACCGGTCAAGTGGGCTGATCAATTAATCTCGCTGGATGACCAAGATGCCA	480		
Qy	481	TTGCTGTGGAAGCTGCCTGCACATAATGTTCCGGCGTTTATTTCTTGATGTCTCTGACACAGA	540		
Db	481	TTGCTGTGGAAGCTGCCTGCACATAATGTTCCGGCGTTTATTTCTTGATGTCTCTGACACAGA	540		
Qy	541	CACCCATCAACAGATTAATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600		
Db	541	CACCCATCAACAGATTAATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600		
Qy	601	TGGTCGCATTTGGGTACACAGCAAAATCGCGCTGTTAGCGGSCCATTAAGTTCCTGCTCGG	660		
Db	601	TGGTCGCATTTGGGTACACAGCAAAATCGCGCTGTTAGCGGSCCATTAAGTTCCTGCTCGG	660		
Qy	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCCGATAG	720		
Db	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCCGATAG	720		
Qy	721	CGGAAACGGGAAAGCGAGCTGGAGTGCCATGTCGCGTTTTTCAACAAACCATGCAAAATGCTGA	780		
Db	721	CGGAAACGGGAAAGCGAGCTGGAGTGCCATGTCGCGTTTTTCAACAAACCATGCAAAATGCTGA	780		
Qy	781	ATGAGGGCATCGTTCACATGTCGATGCTGGTTCGCAACGATCAAGATGGCGCTGGGCGCAA	840		
Db	781	ATGAGGGCATCGTTCACATGTCGATGCTGGTTCGCAACGATCAAGATGGCGCTGGGCGCAA	840		
Qy	841	TGGCGCCATTTACCGAGTCCGGCTGGTTCGGTGGATATCTCGGTAGTGGGATACG	900		
Db	841	TGGCGCCATTTACCGAGTCCGGCTGGTTCGGTGGATATCTCGGTAGTGGGATACG	900		
Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATCAACAGGATTTTC	960		
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATCAACAGGATTTTC	960		
Qy	961	GCCTGCTGGGCAAAACAGCGGTGGACCGCTTGCTGCAAACTCTCTCAAGGCGCAGCGGTGA	1020		

Db 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTCTGCTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAACACACCTTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAACACACCTTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCTCTCCCGCGGTGGCGGATTCATTAAATGCACTGGCAGCACAGGTTT 1140
Db 1081 CGCAACCGCTCTCCCGCGGTGGCGGATTCATTAAATGCACTGGCAGCACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTAGTAGTCACTCAATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTAGTAGTCACTCAATTAG 1200
Qy 1201 GCACAATCTCATGTTTGACAGCTTATCATGACCTGACAGCGTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAATCTCATGTTTGACAGCTTATCATGACCTGACAGCGTGACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGACGCATCGGAAGCTGTGGTATGCTGTGCAAGTGTGTAATCATCTGTAATTCG 1320
Db 1261 TCAGGACGCATCGGAAGCTGTGGTATGCTGTGCAAGTGTGTAATCATCTGTAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACCTCCCGTCTGGAATGATGTTTTCGCGCGACATCATACCGGTT 1380
Db 1321 TGTGCTCAAGCGCACCTCCCGTCTGGAATGATGTTTTCGCGCGACATCATACCGGTT 1380
Qy 1381 CTGGCAATATTTCTGAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATCACTGATA 1440
Db 1381 CTGGCAATATTTCTGAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATCACTGATA 1440
Qy 1441 ATTGTAGCGGATTAACAATTTCAACAGAAACAGCCAGTCCGTTTGGTATGTTTCAAGA 1500
Db 1441 ATTGTAGCGGATTAACAATTTCAACAGAAACAGCCAGTCCGTTTGGTATGTTTCAAGA 1500
Qy 1501 GCACCTTCAACAGGACCATAGATTAATGAACCTGAAGAGGTAACTGGTAACTCGG 1560
Db 1501 GCACCTTCAACAGGACCATAGATTAATGAACCTGAAGAGGTAACTGGTAACTCGG 1560
Qy 1561 ATTAACGCGATAAAGGCTATAACCGTCTCGCTGAAAGTCCGTTAAGAAATTCGAGAAAG 1620
Db 1561 ATTAACGCGATAAAGGCTATAACCGTCTCGCTGAAAGTCCGTTAAGAAATTCGAGAAAG 1620
Qy 1621 ACCGGAATTAAGTCAACGTTGAGGATCCGGAATTAACCTGGAAGAGAAATTCGCAAG 1680
Db 1621 ACCGGAATTAAGTCAACGTTGAGGATCCGGAATTAACCTGGAAGAGAAATTCGCAAG 1680
Qy 1681 CGCGCACTGGCGATGGCCCTGACATATCTTCTGGGCACACGCGCTTGGTGGCTAC 1740
Db 1681 CGCGCACTGGCGATGGCCCTGACATATCTTCTGGGCACACGCGCTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGGACAAAGGTTCCAGGACAAGCTGAT 1800
Db 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGGACAAAGGTTCCAGGACAAGCTGAT 1800
Qy 1801 CCGTTTACCTGGGATGGGTACGTTACAAACGGAAGCTGATGCTTACCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGGGTACGTTACAAACGGAAGCTGATGCTTACCGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCGCAACCGCCCAAAACCTCGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCGCAACCGCCCAAAACCTCGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTTGGCCGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTTGGCCGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGTGGATACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGTGGATACGCTGGCGGAAAGCG 2100

Qy 2101 GGTCTGACCTTCTCGTGGTTGACCTGATTAAAAACAAACACATGATGACAGACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTGGTTGACCTGATTAAAAACAAACACATGATGACAGACCGGATTAC 2160
Qy 2161 TCCATCGCAGAAGCTGCTTTTAAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAAGCTGCTTTTAAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCCAAACATCGACACAGCAAAAGTGAAATTAATGGTAAACGGTACTGCCAGCTTC 2280
Db 2221 GCATGGTCCAAACATCGACACAGCAAAAGTGAAATTAATGGTAAACGGTACTGCCAGCTTC 2280
Qy 2281 AAGGCTCAACCATCCAAACCGTTGCTGGCGTGTGAGCGCAGGTAATTAACGGCCCGCAGT 2340
Db 2281 AAGGCTCAACCATCCAAACCGTTGCTGGCGTGTGAGCGCAGGTAATTAACGGCCCGCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCTGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCTGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
Qy 2401 GAAGCGTTTAAATAAGCAAAACCGTGGTGGCGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGTTTAAATAAGCAAAACCGTGGTGGCGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGGCGAAAGATCCACGATATTGGCGCCACCATGMAAAACGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATATTGGCGCCACCATGMAAAACGCCAGAAAGGTGAAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCCGCTTTCTGATATGCCGTGCTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCCGCTTTCTGATATGCCGTGCTGCTGCTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACCGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACCGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACAAACAAACAAATAACAAATAACAAATAACAACTCGGATCGAGGAGGATTCAGAAATTC 2700
Db 2641 AACAAACAAACAAATAACAAATAACAAATAACAACTCGGATCGAGGAGGATTCAGAAATTC 2700
Qy 2701 GGATCTCTT 2709
Db 2701 GGATCTCTT 2709

RESULT 9
A91965
LOCUS 7475 bp DNA circular PAT 22-JAN-2000
DEFINITION Sequence 1 from Patent WO9822496.
ACCESSION A91965
VERSION A91965.1 GI:6740811
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 7475)
AUTHORS Attwood M.R. and Hurst, D.N.
TITLE ANTIVIRAL PEPTIDE DERIVATIVES
JOURNAL Patent: WO 9822496-A 1 28-MAY-1998;
HOFFMANN LA ROCHE (CH)
FEATURES
source Location/Qualifiers
1. .7475
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 81.8%; Score 2700; DB 6; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATCGTGCAAAACCTTTTCGCGGTATGCGATGATGATAGCGCCCGGAAGAGA 60

Db 1 ||||| CCGACCAATCGAATGGTGC AAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
Qy 61 GTCAATTGAGGTGCTGAATGTGAACACAGTAACGTTATACGATGTCGCAGAGTATGCCG 120
Db 61 GTCAATTGAGGTGCTGAATGTGAACACAGTAACGTTATACGATGTCGCAGAGTATGCCG 120
Qy 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACAGGCCAGCCAGCTTCTCGGAAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACACAGGCCAGCCAGCTTCTCGGAAAA 180
Qy 181 CGCGGAAAAAGTGAAGCGGATGAGCGGAGCTGAATTACATCCCAACCGCTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGGATGAGCGGAGCTGAATTACATCCCAACCGCTGGCAC 240
Qy 241 AACAACTGGCGGCAAAACAGTGGTGTGATTTGGCGTTTGCCACCTCCAGTCTGGGCCCTGC 300
Db 241 AACAACTGGCGGCAAAACAGTGGTGTGATTTGGCGTTTGCCACCTCCAGTCTGGGCCCTGC 300
Qy 301 AGCGCGCTGCGCAAAATTTGTTCGGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGG 360
Db 301 AGCGCGCTGCGCAAAATTTGTTCGGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGG 360
Qy 361 TGGTGGTGTGATGTTAGAACGAAAGCGCGTCCGAGCCCTGTAAAGCGCGGTGCACAATC 420
Db 361 TGGTGGTGTGATGTTAGAACGAAAGCGCGTCCGAGCCCTGTAAAGCGCGGTGCACAATC 420
Qy 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTCCGCTGGATGACCAAGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTCCGCTGGATGACCAAGATGCCA 480
Qy 481 TTGTGTGGAAGCTGCCGTGCAATAATGTTCCGGCGTTATTTCTGATGTTCTTGACCCAGA 540
Db 481 TTGTGTGGAAGCTGCCGTGCAATAATGTTCCGGCGTTATTTCTGATGTTCTTGACCCAGA 540
Qy 541 CACCCATCAACAGTATTTCTCCCATGAAGACGTTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTCTCCCATGAAGACGTTACGCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTGCCTATGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCGTCTCGG 660
Db 601 TGGTGCCTATGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCGTCTCGG 660
Qy 661 CGCGTCTGCGTCTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGCGACTGGAGTGCCATGTCCCGGTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGCGACTGGAGTGCCATGTCCCGGTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAAGCATGAGTGGCGTGGCGCA 840
Db 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAAGCATGAGTGGCGTGGCGCA 840
Qy 841 TCGCGGCCATTAACGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
Db 841 TCGCGGCCATTAACGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
Qy 901 ACGATACCGGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAATCAAAAGGATTTTC 960
Db 901 ACGATACCGGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAATCAAAAGGATTTTC 960
Qy 961 GCCTGTGGGGCAAAACACGCGTGGACCGCTTCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Db 961 GCCTGTGGGGCAAAACACGCGTGGACCGCTTCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTTGCCCGCTCACTGGTGAAGAAAAACCAACCTGGCGGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTTGCCCGCTCACTGGTGAAGAAAAACCAACCTGGCGGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGGTGGCGGATTCATTAATGAGCTGGCAACAGAGTTT 1140

Db 1081 CGCAAAACCGCTCTCCCGCGGTGGCGGATTCATTAATGAGCTGGCACGACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Qy 1201 GCACAATTCATGTTTGACAGCTTATCATCGACTGCAAGGTGCACCAATGCTTCTGGG 1260
Db 1201 GCACAATTCATGTTTGACAGCTTATCATCGACTGCAAGGTGCACCAATGCTTCTGGG 1260
Qy 1261 TCAGGACCCATCGGAAGCTGTGGTGTGAGGTGTCAGGTGCTGAATCACTGTCATAAATCG 1320
Db 1261 TCAGGACCCATCGGAAGCTGTGGTGTGAGGTGTCAGGTGCTGAATCACTGTCATAAATCG 1320
Qy 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATTAATGTTTTTTCGCCGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATTAATGTTTTTTCGCCGACATCAATACGGTT 1380
Qy 1381 CTGGCAATATCTGGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTGTGA 1440
Db 1381 CTGGCAATATCTGGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTGTGA 1440
Qy 1441 ATTTGAGCGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTGTTTCA 1500
Db 1441 ATTTGAGCGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTGTTTCA 1500
Qy 1501 GCCTTCCACCAACGAGGACCATAGATTAATGAAAACTGAAGAGGTAAACTGGTAATCTGG 1560
Db 1501 GCCTTCCACCAACGAGGACCATAGATTAATGAAAACTGAAGAGGTAAACTGGTAATCTGG 1560
Qy 1561 ATTTAAGCGGATTAAGGCTATAACCGTCTCGTGAAGTCGGTAAGAAAAATTCAGAAAGT 1620
Db 1561 ATTTAAGCGGATTAAGGCTATAACCGTCTCGTGAAGTCGGTAAGAAAAATTCAGAAAGT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAAAATTCACAGGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGCTGGAATTAACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGCTGGAATTAACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTTGCTTACCCGATCCGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTTGCTTACCCGATCCGTT 1860
Qy 1861 GAAAGCTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAAAGCTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCTGGATTAAGAACTGAAGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCTGGATTAAGAACTGAAGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACGCTACTTCCCTGGCGCTGATGCTGCTGACGGGGTTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACGCTACTTCCCTGGCGCTGATGCTGCTGACGGGGTTTATCGGTTCAAG 2040
Qy 2041 TATGAAAAACGCAAGTACGATTAAGAACGCTGGCGGTGGATTAACGCTGGCGGCAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGATTAAGAACGCTGGCGGTGGATTAACGCTGGCGGCAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGGTTGACTGATTAATAAACAACATGAATGTCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGGTTGACTGATTAATAAACAACATGAATGTCAGACACCGATTAC 2160
Qy 2161 TCCATCGGAGAGTGCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGGAGAGTGCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220


```
QY 2221 GCATGTCACATCGACACCAAGTGAATTAATGTTGTAACGGTACTGCCACCTTC 2280
Db |||
QY 2221 GCATGTCACATCGACACCAAGTGAATTAATGTTGTAACGGTACTGCCACCTTC 2280
Db |||
QY 2281 AAGGGTCAACCATCCAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAAACGCCGCCAGT 2340
Db |||
QY 2281 AAGGGTCAACCATCCAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAAACGCCGCCAGT 2340
Db |||
QY 2341 CCGAACAAAGAGCTGGCAAAAGATTCCTCGAAAATATCTGCTGACCTGATGAAGGTCG 2400
Db |||
QY 2341 CCGAACAAAGAGCTGGCAAAAGATTCCTCGAAAATATCTGCTGACCTGATGAAGGTCG 2400
Db |||
QY 2401 GAAGCGTTATAAAGACAAACCGTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db |||
QY 2401 GAAGCGTTATAAAGACAAACCGTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db |||
QY 2461 TTGGCGAAGATCCACGTTATTCGCCCAACCATGGAAGAACGCCAGAAAGGTGAATCATG 2520
Db |||
QY 2461 TTGGCGAAGATCCACGTTATTCGCCCAACCATGGAAGAACGCCAGAAAGGTGAATCATG 2520
Db |||
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db |||
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db |||
QY 2581 GCCAGCGTGTGACAGTGTGATGAAGCCCTGAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db |||
QY 2581 GCCAGCGTGTGACAGTGTGATGAAGCCCTGAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db |||
QY 2641 AACAAACAACAATAAACAATAACAACACCTCGGGATCGAGGGAAGGATTCAGAATTC 2700
Db |||
QY 2641 AACAAACAACAATAAACAATAACAACACCTCGGGATCGAGGGAAGGATTCAGAATTC 2700
Db |||

RESULT 10
AR031992
LOCUS AR031992 7475 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5866684.
ACCESSION AR031992
VERSION AR031992.1 GI:5946281
KEYWORDS .
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS 1 (bases 1 to 7475)
Attwood,M.Richard., Hurst,D.Nigel., Jones,P.Stephen.,
Kay,P.Brittain., Raynham,T.Michael. and Willson,F.Xavier.
TITLE Peptidyl inhibitors of viral proteases
JOURNAL Patent: US 5866684-A 1 02-FEB-1999;
FEATURES Location/Qualifiers
SOURCE 1..7475
/molecule="unassigned"
/organism="unassigned"
/origins="unassigned"

ORIGIN
Query Match 81.8%; Score 2700; DB 6; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGAACAAACCTTTCCCGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db |||
QY 1 CCGACACCATCGAATGGTGAACAAACCTTTCCCGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db |||
QY 61 GTCGAATTCAGGGTGGTGAATGTGAACCAAGTAACTGATATGATGTCGAGATGATGCGG 120
Db |||
QY 61 GTCGAATTCAGGGTGGTGAATGTGAACCAAGTAACTGATATGATGTCGAGATGATGCGG 120
Db |||
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGTAACCAAGCGCAGCGCTTCTCGGAAA 180
Db |||
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGTAACCAAGCGCAGCGCTTCTCGGAAA 180
Db |||
QY 181 CGCGGGAAGATGGAAGCGCGATGCGGAGCTGAATTAATTCACCAACCGCGTGGCAC 240
Db |||
```

```
Db |||
QY 181 CGCGGGAAGATGGAAGCGCGATGCGGAGCTGAATTAATTCACCAACCGCGTGGCAC 240
QY |||
QY 241 AACAACTGGCGGCAAAACAGTCTGTTGTTGATGCGGTGCGACCTCTCAGTCTGCGCTTCG 300
Db |||
QY 241 AACAACTGGCGGCAAAACAGTCTGTTGTTGATGCGGTGCGACCTCTCAGTCTGCGCTTCG 300
Db |||
QY 301 ACGCCGCTGCAAAATTTGTCGGCGGATTAATCTCTCGCGGATCAACTGGGTGCGGACG 360
Db |||
QY 301 ACGCCGCTGCAAAATTTGTCGGCGGATTAATCTCTCGCGGATCAACTGGGTGCGGACG 360
Db |||
QY 361 TGTGTTGTTGATGTTAGAAACGAAGCGCGCTGGAAGCTCTGTAAGCGCGGTGCAAACT 420
Db |||
QY 361 TGTGTTGTTGATGTTAGAAACGAAGCGCGCTGGAAGCTCTGTAAGCGCGGTGCAAACT 420
Db |||
QY 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTATCCGCTGGATGACCAAGATGCCA 480
Db |||
QY 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTATCCGCTGGATGACCAAGATGCCA 480
Db |||
QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCTTATTTCTTGATGTTCTTGACCGAGA 540
Db |||
QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCTTATTTCTTGATGTTCTTGACCGAGA 540
Db |||
QY 541 CACCATCAACAGTATTAATTTCTCCCAAGAACGCTACGCGACTGGCGGTGGAGCATC 600
Db |||
QY 541 CACCATCAACAGTATTAATTTCTCCCAAGAACGCTACGCGACTGGCGGTGGAGCATC 600
Db |||
QY 601 TGTGCGCATTTGGGTTCACCGCAATCGCGCTGTTAGCGGCGCCATTAAGTCTGCTCGG 660
Db |||
QY 601 TGTGCGCATTTGGGTTCACCGCAATCGCGCTGTTAGCGGCGCCATTAAGTCTGCTCGG 660
Db |||
QY 661 CGCGTCTGCGTCTGCGTGGCTGCGCAATAATCTCACTCGCAATCAAAATCAAGCTGCTCG 720
Db |||
QY 661 CGCGTCTGCGTCTGCGTGGCTGCGCAATAATCTCACTCGCAATCAAAATCAAGCTGCTCG 720
Db |||
QY 721 CGGAAACGGGAAGCGCTGAGTGCCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db |||
QY 721 CGGAAACGGGAAGCGCTGAGTGCCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db |||
QY 781 ATGAGGCGATCGTTCCCGCTGCGATGCTGTTGCCAACGATCAGATGGCGCTGGCGCAA 840
Db |||
QY 781 ATGAGGCGATCGTTCCCGCTGCGATGCTGTTGCCAACGATCAGATGGCGCTGGCGCAA 840
Db |||
QY 841 TGCAGCGCATTAACGAGTCCGGCTGCGGTGCGGATATCTCGGTAGTGGGATACG 900
Db |||
QY 841 TGCAGCGCATTAACGAGTCCGGCTGCGGTGCGGATATCTCGGTAGTGGGATACG 900
Db |||
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCTGTAACCAACCATCAAAACAGGATTTTC 960
Db |||
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCTGTAACCAACCATCAAAACAGGATTTTC 960
Db |||
QY 961 GCCTGCTGGGGCAAAACAGCGCTGCGGCTGCGGTGCGGATATCTCGGTAGTGGGATACG 1020
Db |||
QY 961 GCCTGCTGGGGCAAAACAGCGCTGCGGCTGCGGTGCGGATATCTCGGTAGTGGGATACG 1020
Db |||
QY 1021 AGGCAATCAGCTGTTGCGGCTCTCACTGTTGAAAGAAAAACCAACCTCTGGCGCCCAATA 1080
Db |||
QY 1021 AGGCAATCAGCTGTTGCGGCTCTCACTGTTGAAAGAAAAACCAACCTCTGGCGCCCAATA 1080
Db |||
QY 1081 CGCAACCGCTCTCCCGCGGCTGCGGCTGCGGATTAATTAATGAGCTGGGACGAGAGTTT 1140
Db |||
QY 1081 CGCAACCGCTCTCCCGCGGCTGCGGCTGCGGATTAATTAATGAGCTGGGACGAGAGTTT 1140
Db |||
QY 1141 CCGCATGGAACCGCGCAGTGAAGCGCAACGCAATTAATGAGCTGGGACGAGAGTTT 1200
Db |||
QY 1141 CCGCATGGAACCGCGCAGTGAAGCGCAACGCAATTAATGAGCTGGGACGAGAGTTT 1200
Db |||
QY 1201 GCACAACTCTCATGTTGACAGCTTATCATGCTGCAAGGTGCAACAAATGCTTCTGGCG 1260
Db |||
QY 1201 GCACAACTCTCATGTTGACAGCTTATCATGCTGCAAGGTGCAACAAATGCTTCTGGCG 1260
Db |||
QY 1261 TCAGGCGAGCATCGGAAGCTGTTGCTGCTGCGAGGTGCTGCAATCACTGCAATTAATTCG 1320
Db |||
QY 1261 TCAGGCGAGCATCGGAAGCTGTTGCTGCTGCGAGGTGCTGCAATCACTGCAATTAATTCG 1320
Db |||
```

QY 1321 TGTGCTCAAGCGGCACCTCCCGTTCTGGATATGTTTTTGGCCGACATCATACGGTT 1380
DB 1321 TGTGCTCAAGCGGCACCTCCCGTTCTGGATATGTTTTTGGCCGACATCATACGGTT 1380
QY 1381 CTGGCAAAATATTCTGAAATAGCTGTGACAAATTAATCATCGGCTCGTATAATGTTGTGA 1440
DB 1381 CTGGCAAAATATTCTGAAATAGCTGTGACAAATTAATCATCGGCTCGTATAATGTTGTGA 1440
QY 1441 ATTGTAGCGGATACAAATTTTACACAGAAACAGCCAGTCGGTTTAGGTTTTCAGGA 1500
DB 1441 ATTGTAGCGGATACAAATTTTACACAGAAACAGCCAGTCGGTTTAGGTTTTCAGGA 1500
QY 1501 GCATCTCACACACAGGACCATAGATTATGAAACTGAAGAGTAAACTGGTATCTGG 1560
DB 1501 GCATCTCACACACAGGACCATAGATTATGAAACTGAAGAGTAAACTGGTATCTGG 1560
QY 1561 ATTTAAGCGCGATAAAGGCTATAACCGTCTCGCTGAAATCGGTAAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTTAAGCGCGATAAAGGCTATAACCGTCTCGCTGAAATCGGTAAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCACCGTTGAGCATCCGGATAAATTCGAGAGATTTCCACAGGTT 1680
DB 1621 ACCGGAATTAAGTCACCGTTGAGCATCCGGATAAATTCGAGAGATTTCCACAGGTT 1680
QY 1681 GCGCAACTGGCGATGGCGCTGACATTTATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
DB 1681 GCGCAACTGGCGATGGCGCTGACATTTATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGGCTGTGGCTGAAATCACCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
DB 1741 GCTCAATCTGGGCTGTGGCTGAAATCACCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
QY 1801 CGTTTACCTGGGATGCGGTACGTTTACAAAGCGCAAGCTGATGCTTACCCGATCGCTGT 1860
DB 1801 CGTTTACCTGGGATGCGGTACGTTTACAAAGCGCAAGCTGATGCTTACCCGATCGCTGT 1860
QY 1861 GAAGGCTTATCGCTGATTTAACAAGATCTGCTGCGGACCCCGCCAAAACCTTGGGAA 1920
DB 1861 GAAGGCTTATCGCTGATTTAACAAGATCTGCTGCGGACCCCGCCAAAACCTTGGGAA 1920
QY 1921 GAGATCCCGGCGCTCGATTAAGAACTGAAGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
DB 1921 GAGATCCCGGCGCTCGATTAAGAACTGAAGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
DB 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGGCAAGTACGACATTAAGACGTGGGGTGGATAACGCTGGCGGAAAGCG 2100
DB 2041 TATGAAACCGGCAAGTACGACATTAAGACGTGGGGTGGATAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGTTGACCTGATTAAGAAACACATCAATGACCATCAAGCGGCGCTGG 2160
DB 2101 GGTCTGACCTTCTGTTGACCTGATTAAGAAACACATCAATGACCATCAAGCGGCGCTGG 2160
QY 2161 TCCATCGCAAGAGCTGCGCTTTTAAAGCGGAAAACAGCATCAACATCAAGCGGCGCTGG 2220
DB 2161 TCCATCGCAAGAGCTGCGCTTTTAAAGCGGAAAACAGCATCAACATCAAGCGGCGCTGG 2220
QY 2221 GCATGGTCCAAATCGACACAGCAAGAGTGAATTTATGGTGAACGGTATGCGGACCTTC 2280
DB 2221 GCATGGTCCAAATCGACACAGCAAGAGTGAATTTATGGTGAACGGTATGCGGACCTTC 2280
QY 2281 AAGGCTCAACATCCAAACCGTTCGTTGGCGTGTGCTGAGCGAGGATTTAAGCGCCCGAT 2340
DB 2281 AAGGCTCAACATCCAAACCGTTCGTTGGCGTGTGCTGAGCGAGGATTTAAGCGCCCGAT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGCTG 2400
DB 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGCTG 2400

QY 2401 GAAGCGTTAATAAGACAAACCCCTGGGTGCGCTAGCGCTGAAGCTTTACGAGGAAG 2460
DB 2401 GAAGCGTTAATAAGACAAACCCCTGGGTGCGCTAGCGCTGAAGCTTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGCTATTTCGCCGCCACCATGGAACACGCCAGAAAGGTGAAATCATG 2520
DB 2461 TTGGCGAAAGATCCACGCTATTTCGCCGCCACCATGGAACACGCCAGAAAGGTGAAATCATG 2520
QY 2521 CCGAACATCCCGAGATGTCGGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAACGCC 2580
DB 2521 CCGAACATCCCGAGATGTCGGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
DB 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAATAACAATAACAACCAACCTCGGATCGAGGGAAGGATTTTCAGAAATTC 2700
DB 2641 AACAAACAAACAATAACAATAACAACCAACCTCGGATCGAGGGAAGGATTTTCAGAAATTC 2700

RESULT 11
AR207294
LOCUS AR207294 7475 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 1 from patent US 6372883.
ACCESSION AR207294
VERSION AR207294.1 GI:21506162
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7475)
AUTHORS Attwood,M.Richard., Hurst,D.Nigel., Jones,P.Stephen.,
Kay,P.Brittain., Raynham,T.Michael. and Wilson,F.Xavier.
TITLE Antiviral medicaments
JOURNAL Patent: US 6372883-A 1 16-APR-2002;
FEATURES Location/Qualifiers
source 1..7475
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 81.8%; Score 2700; DB 6; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
DB 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTCCAGGTGGTGAATGTGAAACCAAGTAAAGTTATACGATGTCGACAGTATGCG 120
DB 61 GTCAATTCCAGGTGGTGAATGTGAAACCAAGTAAAGTTATACGATGTCGACAGTATGCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGCGCCAGCCACGCTTTTCGGA 180
DB 121 GTGCTCTTATCAGACCGTTTCCCGCGTGGTGAACCAAGCGCCAGCCACGCTTTTCGGA 180
QY 181 CGCGGAAAAGTGAAGCGCGATGGCGGAGCTGAATTTACATTTCCCAACCGCTGGGCAC 240
DB 181 CGCGGAAAAGTGAAGCGCGATGGCGGAGCTGAATTTACATTTCCCAACCGCTGGGCAC 240
QY 241 AACAACTGGCGGCAAAACAGTCTGTTGCTGATTTGGCGTTGCCACCTTCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGCAAAACAGTCTGTTGCTGATTTGGCGTTGCCACCTTCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTTCGCAATTTGTCGCGGGATTAATCTCGCGCGATCAACTGGGTGCGAGG 360
DB 301 ACGCGCGTTCGCAATTTGTCGCGGGATTAATCTCGCGCGATCAACTGGGTGCGAGG 360
QY 361 TGGTGGTCTCGATGTAAGACGAGCGGCTGAGAGCTGTAAAGCGCGGTGCAACAATC 420
DB 361 TGGTGGTCTCGATGTAAGACGAGCGGCTGAGAGCTGTAAAGCGCGGTGCAACAATC 420

QY 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACATATCCGCTGGATACACGAGGATGCCA 480
DB |||||
DB 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAACATATCCGCTGGATACACGAGGATGCCA 480
QY 481 TTGCTGTGGAGAGCTGCCCTGCACTAATGTTCGGCGTTTATTTCTTGATGCTCTGACACAGA 540
DB |||||
DB 481 TTGCTGTGGAGAGCTGCCCTGCACTAATGTTCGGCGTTTATTTCTTGATGCTCTGACACAGA 540
QY 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGCGACTGGCGCTGGAGCATC 600
DB |||||
DB 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGCGACTGGCGCTGGAGCATC 600
QY 601 TGGTCGCAATGGGTACACAGCAAAATCGGCGTGTATAGCGGGGCCCAATTAAGTTCTGTCTCGG 660
DB |||||
DB 601 TGGTCGCAATGGGTACACAGCAAAATCGGCGTGTATAGCGGGGCCCAATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGGCTGCGCTGGCTGGGCAATAATATCTCACTCGCAATCAAAATTCAGCGCATAG 720
DB |||||
DB 661 CGCGTCTCGGCTGCGCTGGCTGGGCAATAATATCTCACTCGCAATCAAAATTCAGCGCATAG 720
QY 721 CGGAACGGGAAGGACATGGAGTGCATGTCCGGTTTTCAACAACCAATGCAAAATGCTGA 780
DB |||||
DB 721 CGGAACGGGAAGGACATGGAGTGCATGTCCGGTTTTCAACAACCAATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTCGATGCTGGTTGCCAACGATCAGATGGCGCTGGCGCGCA 840
DB |||||
DB 781 ATGAGGGCATCGTTCCCACTCGATGCTGGTTGCCAACGATCAGATGGCGCTGGCGCGCA 840
QY 841 TGGCGCCATTAACGAGTCCGGGCTGCGGTGGTGGGATATCTCGGTAGTGGGATACG 900
DB |||||
DB 841 TGGCGCCATTAACGAGTCCGGGCTGCGGTGGTGGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAAACCAACATCAAAACAGGATTTTC 960
DB |||||
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAAACCAACATCAAAACAGGATTTTC 960
QY 961 GCCTGTGGGGCAACACAGCGTGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
DB |||||
DB 961 GCCTGTGGGGCAACACAGCGTGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
QY 1021 AGGGCAATACGCTGTGCGCGCTCAGTGGTGAAGAAAGAAACACACCTGGCGCGCAATA 1080
DB |||||
DB 1021 AGGGCAATACGCTGTGCGCGCTCAGTGGTGAAGAAAGAAACACACCTGGCGCGCAATA 1080
QY 1081 CGCAAAACGCGCTCTCCGCGCGCTTGGCGGATTCATTAATGACGCTGGCAGCAGGTTT 1140
DB |||||
DB 1081 CGCAAAACGCGCTCTCCGCGCGCTTGGCGGATTCATTAATGACGCTGGCAGCAGGTTT 1140
QY 1141 CCGGACTGGAAGCGGCGAGTGAGGCAACGCAATTAATGAGTTAGTCACTCATTTAG 1200
DB |||||
DB 1141 CCGGACTGGAAGCGGCGAGTGAGGCAACGCAATTAATGAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAGGTGCACCAATGCTTCTCGCG 1260
DB |||||
DB 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAGGTGCACCAATGCTTCTCGCG 1260
QY 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGACGCTGCAATCACTGCATAATTCG 1320
DB |||||
DB 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGACGCTGCAATCACTGCATAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACTCCGGTTCTGGATAATGTTTTTTCGCGCGACATCAAAACGGTT 1380
DB |||||
DB 1321 TGTGCTCAAGCGGCACTCCGGTTCTGGATAATGTTTTTTCGCGCGACATCAAAACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
DB |||||
DB 1381 CTGGCAAAATTTCTGAAATAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
QY 1441 ATTGTAGCGGATCAAAATTTTACAGGAAACAGCGAGTTCGGTTTACCGA 1500
DB |||||
DB 1441 ATTGTAGCGGATCAAAATTTTACAGGAAACAGCGAGTTCGGTTTACCGA 1500

QY 1501 GCACCTTCAACAACGAGCATATGATTATGAATACTGAAGAGGTAACTGGTAACTCG 1560
DB |||||
DB 1501 GCACCTTCAACAACGAGCATATGATTATGAATACTGAAGAGGTAACTGGTAACTCG 1560
QY 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAGTCCGTTAAGAAATTCGAGAAAGAT 1620
DB |||||
DB 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAGTCCGTTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTACCGTTGAGCATCCGATAACTGGAAGAGAAATTCACACAGTT 1680
DB |||||
DB 1621 ACCGGAATTAAGTACCGTTGAGCATCCGATAACTGGAAGAGAAATTCACACAGTT 1680
QY 1681 GCGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTGTGGGCTAC 1740
DB |||||
DB 1681 GCGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTGTGGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTTGGTGAATAACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
DB |||||
DB 1741 GCTCAATCTGGCTGTTGGTGAATAACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTACACGGCAAGCTGATTGTTACCGGATCGCTGTT 1860
DB |||||
DB 1801 CCGTTTACCTGGGATGCGGTACGTTACACGGCAAGCTGATTGTTACCGGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTCGGAA 1920
DB |||||
DB 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTCGGAA 1920
QY 1921 GAGATCCGCGGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
DB |||||
DB 1921 GAGATCCGCGGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTCAGAAACCGTACCTTACCTCGCGCTGATTGCTGCTGACGGGGTTATGGTTCAAG 2040
DB |||||
DB 1981 CTCAGAAACCGTACCTTACCTCGCGCTGATTGCTGCTGACGGGGTTATGGTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTAAGAGCGTGGCGTGGATAACGCTGGCGGAAAGCG 2100
DB |||||
DB 2041 TATGAAACCGCAAGTACGACATTAAGAGCGTGGCGTGGATAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGTTGACCTGATTAAACAAACACATGAATGCGACACACCGATTAC 2160
DB |||||
DB 2101 GGTCTGACCTTCTGTTGACCTGATTAAACAAACACATGAATGCGACACACCGATTAC 2160
QY 2161 TCCATCGCAAGCTGCTCTTTAATAAGGCGAAACAGCGATGACCATCAACGCGCGGTGG 2220
DB |||||
DB 2161 TCCATCGCAAGCTGCTCTTTAATAAGGCGAAACAGCGATGACCATCAACGCGCGGTGG 2220
QY 2221 GCATGTTCCAAATCGACACACAGCAAGTGAATTTATGGTGTAAACGCTGCTGCGACCTTC 2280
DB |||||
DB 2221 GCATGTTCCAAATCGACACACAGCAAGTGAATTTATGGTGTAAACGCTGCTGCGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTGCGTGGCGTCTGAGCGCAGGTATTAAACGCGCGAGT 2340
DB |||||
DB 2281 AAGGGTCAACCATCCAAACCGTTGCGTGGCGTCTGAGCGCAGGTATTAAACGCGCGAGT 2340
QY 2341 CCGAAACAAAGAGCTGCGCAAAAGAGTTCTCGTGAATACTATCTGCTGACTGATGAAGGTCTG 2400
DB |||||
DB 2341 CCGAAACAAAGAGCTGCGCAAAAGAGTTCTCGTGAATACTATCTGCTGACTGATGAAGGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAGAG 2460
DB |||||
DB 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAGAG 2460
QY 2461 TTGGCGGAAAGAGTCCAGTATGCGCGCACCATGGAAACCGCCAGAAAGGTGAATCATG 2520
DB |||||
DB 2461 TTGGCGGAAAGAGTCCAGTATGCGCGCACCATGGAAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCGTTTCTGGTATGCGGTGCGTACTCGCGGTGATCAACGCC 2580
DB |||||
DB 2521 CCGAACATCCCGCAGATGTCGCGTTTCTGGTATGCGGTGCGTACTCGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGGTGCTCAGACTGTCGATGAGGCCCTGAAAGACCGCGCAGACTAATTCGAGCTCG 2640

Db	2581	GCCAGCGGTGCTCAGATGTCGATGTAAGCCCTGAAAGCGCGCAGACTAAATTCGAGCTCG	2640
Qy	2641	AACAAACAACAATAACAATAACAACAACTCGGGATCGAGGGAAGATTTTCAGAAATTC	2700
Db	2641	AACAAACAACAATAACAATAACAACAACTCGGGATCGAGGGAAGATTTTCAGAAATTC	2700
RESULT 12			
AX377532			
LOCUS	AX377532	9191 bp	DNA linear PAT 18-MAR-2002
DEFINITION	Sequence 9 from Patent W00212553.		
ACCESSION	AX377532		
VERSION	AX377532.1	GI:19573718	
KEYWORDS			
SOURCE	synthetic construct		
ORGANISM	other sequences; artificial sequences.		
REFERENCE	1		
AUTHORS	Kappel,A., Polakowski,T., Pignot,M., Windhab,N., Behrensdozf,H. and Muth,J.		
TITLE	Method for detecting mutations in nucleotide sequences		
JOURNAL	Patent: WO 0212553-A 9 14-FEB-2002;		
FEATURES	Nanogen Recognomics GmbH (DE)		
source	Location/Qualifiers		
	1..9191		
	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	/note="pMALc2x-mutS Plasmid"		
ORIGIN			
Query Match	81.8%;	Score 2698;	DB 6; Length 9191;
Best Local Similarity	99.8%;	Pred. No. 0;	
Matches 2701;	Conservative 0;	Mismatches 5;	Indels 0; Gaps 0;
Qy	1	CCGACACCATCGAATGTGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA	60
Db	1	CCGACACCATCGAATGTGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA	60
Qy	61	GTCAATTCAGGGTGGTGAATGTGAACCAAGTAAAGTTATACGATGTCGCAGAGTATCCG	120
Db	61	GTCAATTCAGGGTGGTGAATGTGAACCAAGTAAAGTTATACGATGTCGCAGAGTATCCG	120
Qy	121	GTGTCCTTATCAGACCGCTTTCCCGCGTGGTGAACCAAGCCAGCCAGCCAGCTTTCTCGCAAAA	180
Db	121	GTGTCCTTATCAGACCGCTTTCCCGCGTGGTGAACCAAGCCAGCCAGCCAGCTTTCTCGCAAAA	180
Qy	181	CGCGGAAAAAGTGGAAAGCGCGATGCGGAGCTGAATTAATTCACAAACCGGTGGCAC	240
Db	181	CGCGGAAAAAGTGGAAAGCGCGATGCGGAGCTGAATTAATTCACAAACCGGTGGCAC	240
Qy	241	AACAACTGGCGGGCAAAACAGTCTGTTGATGTGGCTTGCACCTCCAGTCTGGCCCTGC	300
Db	241	AACAACTGGCGGGCAAAACAGTCTGTTGATGTGGCTTGCACCTCCAGTCTGGCCCTGC	300
Qy	301	ACGCGCGCTCGCAAAATGTCCGCGCATTTAAATCTCCGCGCGATCAACTGGGTGCCAGCG	360
Db	301	ACGCGCGCTCGCAAAATGTCCGCGCATTTAAATCTCCGCGCGATCAACTGGGTGCCAGCG	360
Qy	361	TGTTGGTGTGATGATGAAGCAAGCGCGCTGCGAAGCTGTGAAAGCGCGGTGCACAATC	420
Db	361	TGTTGGTGTGATGATGAAGCAAGCGCGCTGCGAAGCTGTGAAAGCGCGGTGCACAATC	420
Qy	421	TTCTCGGCACCGGTCAAGTGGGCTGATCATTAATCTCCGCTGGATGACCAAGATGCCA	480
Db	421	TTCTCGGCACCGGTCAAGTGGGCTGATCATTAATCTCCGCTGGATGACCAAGATGCCA	480
Qy	481	TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTCTCACCAGA	540
Db	481	TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTCTCACCAGA	540
Qy	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGA CGGTACGCGACTGGGCGTGGAGCATC	600

Db	1621	ACCGAAATAAAGTCACCGTTGAGCATCCGGATAAATGGAGAGAAATTTCCACAGGTT	1680
Qy	1681	GCGGCAACTGCGAGTGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
Db	1681	GCGGCAACTGCGAGTGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
Qy	1741	GCTCAATCTGGCCTGTTGGCTGAAATCAACCCGGACAAAGCGTTCAGGACAAGCTGTAT	1800
Db	1741	GCTCAATCTGGCCTGTTGGCTGAAATCAACCCGGACAAAGCGTTCAGGACAAGCTGTAT	1800
Qy	1801	CGGTTTACTCGGATGCGGTACGTTTACAAACGGCAAGCTGATGCTTACCGATCGCTGTT	1860
Db	1801	CGGTTTACTCGGATGCGGTACGTTTACAAACGGCAAGCTGATGCTTACCGATCGCTGTT	1860
Qy	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCAAAACCTGGCAA	1920
Db	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCAAAACCTGGCAA	1920
Qy	1921	GAGATCCGGCGCTGGATTAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCGGCGCTGGATTAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
Qy	1981	CTGCAAGAACCGTACTTACACTGCGCGCTGATTGCTGCTGACGGGGGTATGCGTTCAG	2040
Db	1981	CTGCAAGAACCGTACTTACACTGCGCGCTGATTGCTGCTGACGGGGGTATGCGTTCAG	2040
Qy	2041	TATGAAACCGCAAGTAGCATTAAAGACGTGGCGCTGGATTAACGTGCGCGGAAGCG	2100
Db	2041	TATGAAACCGCAAGTAGCATTAAAGACGTGGCGCTGGATTAACGTGCGCGGAAGCG	2100
Qy	2101	GGTCTGACCTTCCTGGTTGACCTGATTTAAAAACAAACACATGATGAGACACCGATTAC	2160
Db	2101	GGTCTGACCTTCCTGGTTGACCTGATTTAAAAACAAACATGATGAGACACCGATTAC	2160
Qy	2161	TCCATCCGAGAAGCTGCTTTATATAAGGCGAAACACGCGATGACCATCAACGCCCGGTG	2220
Db	2161	TCCATCCGAGAAGCTGCTTTATATAAGGCGAAACACGCGATGACCATCAACGCCCGGTG	2220
Qy	2221	GCATGTTCCAACATCGACACCGCAAGTCAATTTATGGTGTACCGTACTGCGCGACCTTC	2280
Db	2221	GCATGTTCCAACATCGACACCGCAAGTCAATTTATGGTGTACCGTACTGCGCGACCTTC	2280
Qy	2281	AAGGTCACCATCCAAACCGTTTCGTTGGCGTGTGAGCGCAGGTATTAACGCCGCAGT	2340
Db	2281	AAGGTCACCATCCAAACCGTTTCGTTGGCGTGTGAGCGCAGGTATTAACGCCGCAGT	2340
Qy	2341	CCGAAACAAGAGCTGGCAAAAGATTCTCGAAAACTATCTGCTGACTGATGAAGTCTG	2400
Db	2341	CCGAAACAAGAGCTGGCAAAAGATTCTCGAAAACTATCTGCTGACTGATGAAGTCTG	2400
Qy	2401	GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG	2460
Db	2401	GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG	2460
Qy	2461	TTGGCGAAGATCGACGTATTTGCGGCCACATGGAAAAACGCCAGAAAAAGTGAATCATG	2520
Db	2461	TTGGCGAAGATCGACGTATTTGCGGCCACATGGAAAAACGCCAGAAAAAGTGAATCATG	2520
Qy	2521	CCGAAACATCCGCGAGATGTCGCTTTCTGTTATGCGGTGCGTACTGCGGTGATCAACGCC	2580
Db	2521	CCGAAACATCCGCGAGATGTCGCTTTCTGTTATGCGGTGCGTACTGCGGTGATCAACGCC	2580
Qy	2581	GCCAGCGGTGTCGACTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
Db	2581	GCCAGCGGTGTCGACTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
Qy	2641	AACAACAACAACATTAACATTAACACACCTCGGGATCGAGGGAAGGATTTTCAGAAATC	2700
Db	2641	AACAACAACAACATTAACATTAACACACCTCGGGATCGAGGGAAGGATTTTCAGAAATC	2700
Qy	2701	GGATCC	2706
Db	2701	GGATCC	2706

RESULT 13	AF097412	6724 bp	DNA	circular SYN 28-JAN-2002
LOCUS	AF097412	6724 bp	DNA	circular SYN 28-JAN-2002
DEFINITION	AF097412	6724 bp	DNA	circular SYN 28-JAN-2002
ACCESSION	AF097412	6724 bp	DNA	circular SYN 28-JAN-2002
VERSION	AF097412.1	GI:3983120		
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
MEDLINE				
PUBMED				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
FEATURES				
source				
gene				
CDS				
promoter				
gene				
CDS				
misc_feature				
misc_feature				
misc_feature				

gene /note="multiple cloning site"
3426..4356
/gene="bla"
promoter /gene="bla"
3426..3454
/gene="bla"
CDS /gene="bla"
3496..4356
/gene="bla"
/codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="AAC83814.1"
/db_xref="GI:3983123"
/translation="MSIOHFRVALIPPEAAFCPLVFAPHETILVKVDAEDOLGARVGY
IELDNLNSKILLESFPEPERSPMSTFKVLLCGAVLSRVDGBOELGRIIHYSONDLVE
YSPVTEKHLTDGMTVELCSRAITMSDNTAANLLTTTGGPKELTFLRHNGDHDVRL
DRWPELNEAI PNDRDTPMPVAMATTLKLLTGSLTLASRQQLIDIMEADKVAGPL
LRSLAPAGWFIADKSGAGERSGRIIAALGPDGKPSRIWVIYTTGSQATMDERNQIA
BIGASLIXHW"

ORIGIN

Query Match 81.1%; Score 2676.2; DB 12; Length 6724;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2678; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGCBAACCTTTCCGGGTATGGCATAGAGCCCGGAAGAGA 60
DB 1 CCGACACCATCGAATGGTGCBAACCTTTCCGGGTATGGCATAGAGCCCGGAAGAGA 60
QY 61 GTCAATTCAAGGTGGTGAATGTGAACCAAGTAACTTATACCATGTTCGACAGTATGCG 120
DB 61 GTCAATTCAAGGTGGTGAATGTGAACCAAGTAACTTATACCATGTTCGACAGTATGCG 120
QY 121 GTGTCCTTATCAGACCGGTTTCCCGGTGGTGAACCAAGCCAGCCACGTTTCTGGAAAA 180
DB 121 GTGTCCTTATCAGACCGGTTTCCCGGTGGTGAACCAAGCCAGCCAGTTTCTGGAAAA 180
QY 181 CGCGGAAAAAGTGAAGCGCGCATGGCGAGCTGAATTAATCCCAACCGGTGGCAC 240
DB 181 CGCGGAAAAAGTGAAGCGCGCATGGCGAGCTGAATTAATCCCAACCGGTGGCAC 240
QY 241 AACAACTGGCGGCAAAAGTGTGCTGATTGGGGTTGCCACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGCAAAAGTGTGCTGATTGGGGTTGGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTTCGCAAAATGTCGGCGCATTAATCTCGCGCGCATCAACTGGGTGGCAGCG 360
DB 301 ACGCGCGTTCGCAAAATGTCGGCGCATTAATCTCGCGCGCATCAACTGGGTGGCAGCG 360
QY 361 TGGTGGTTCGATGTGTAGAACGAAGCGCGTCGAAGCCTGTAAAGCGGCGTGCACAATC 420
DB 361 TGGTGGTTCGATGTGTAGAACGAAGCGCGTCGAAGCCTGTAAAGCGGCGTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTCCGCTGGATGACACAGGATGCCA 480
DB 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTCCGCTGGATGACACAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCAATAATGTTCCGGCGTTATTTCTGATGTCTCTGACCAGA 540
DB 481 TTGCTGTGGAAGCTGCTGCAATAATGTTCCGGCGTTATTTCTGATGTCTCTGACCAGA 540
QY 541 CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGCGATCGGCGTGGAGCATC 600
DB 541 CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGCGATCGGCGTGGAGCATC 600
QY 601 TGGTGCATTTGGGTTCACAGCAAAATCGCGCTGTAGCGGCGCCATTAAGTCTGTCTCGG 660
DB 601 TGGTGCATTTGGGTTCACAGCAAAATCGCGCTGTAGCGGCGCCATTAAGTCTGTCTCGG 660
QY 661 CGCGTCTGGCTGTGGCTGGCATTAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGGCTGTGGCTGGCATTAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGGGACCTGGAGTGCAATGTCGGGTTTTCACAAACCATGCAAAATGCTGA 780

DB 721 CGGAACGGGAAGGGGACCTGGAGTGCCATGTCGGGTTTTCACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTGCGATGCTGGTGGCAACGATCAGATGCGCTGGGCGCAA 840
DB 781 ATGAGGGCATCGTTCCCACTGCGATGCTGGTGGCAACGATCAGATGCGCTGGGCGCAA 840
QY 841 TGGCGCCATTACCGAGTCGGGCTGGCGTGGTGGCGGATATCTCGTAGTGGGATAG 900
DB 841 TGGCGCCATTACCGAGTCGGGCTGGCGTGGTGGCGGATATCTCGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATCAAAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAAAACCAAGCGTGACCGCTTGTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
DB 961 GCCTGCTGGGCAAAACCAAGCGTGACCGCTTGTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCGTCTCAGCTGGTGAAGAAAAAACCAACCTGGCGCCCAATA 1080
DB 1021 AGGGCAATCAGCTGTTGCCCGTCTCAGCTGGTGAAGAAAAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGCAGCTGCACACAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGCAGCTGCACACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
DB 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAAATTCATGTTTGACAGCTTATCATCGACGGTGACCAATGCTTCTGGCG 1260
DB 1201 GCACAAATTCATGTTTGACAGCTTATCATCGACGGTGACCAATGCTTCTGGCG 1260
QY 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTGAATCACTGCATTAATTCG 1320
DB 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTGAATCACTGCATTAATTCG 1320
QY 1321 TGTGCTCAAGCGCACCTCCCGTCTCGGATTAATGTTTTTGGCGCGACATCAATACGGTT 1380
DB 1321 TGTGCTCAAGCGCACCTCCCGTCTCGGATTAATGTTTTTGGCGCGACATCAATACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTCACAAATTAATCATCGGCTCGTATATGTTGGA 1440
DB 1381 CTGGCAAAATTTCTGAAATGAGCTGTTCACAAATTAATCATCGGCTCGTATATGTTGGA 1440
QY 1441 ATTGTAGCGGATTAACAAATTTTCACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500
DB 1441 ATTGTAGCGGATTAACAAATTTTCACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACTGAAGAAGGTAATACTGGTAATCTGG 1560
DB 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACTGAAGAAGGTAATACTGGTAATCTGG 1560
QY 1561 ATTAACCGCGATAAAGGCTATAACCGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTAACCGCGATAAAGGCTATAACCGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAAAATCCCAAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAAAATCCCAAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACAGCCGCTTTGGTGGCTAC 1740
DB 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACAGCCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCAATCCCGGCAAAAGCGTTCCAGGACAAGCTGAT 1800
DB 1741 GCTCAATCTGGCTGTGGCTGAAATCAATCCCGGCAAAAGCGTTCCAGGACAAGCTGAT 1800
QY 1801 CCGTTTACCTGGATGCGGTAGCTTTACAAAGCGCAAGCTGATGCTTACCCGATCCCTGTT 1860

Db 841 TCGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGGATATCTCGGTAGTGGGATACG 900

Qy 901 ACGATACCGAAGACAGCTCATGTTTATATCCGCGCTTAAACCAACCATCAAAACAGGATTTTC 960

Db 901 ACGATACCGAAGACAGCTCATGTTTATATCCGCGCTTAAACCAACCATCAAAACAGGATTTTC 960

Qy 961 GCCTGCTGGGCGAACCAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020

Db 961 GCCTGCTGGGCGAACCAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020

Qy 1021 AGGCGAATCAGCTGTTGGCCGCTCTCACTGGTGAAGAAACCAACCTGGCGGCCAATA 1080

Db 1021 AGGCGAATCAGCTGTTGGCCGCTCTCACTGGTGAAGAAACCAACCTGGCGGCCAATA 1080

Qy 1081 CGCAAAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGCGAGCTGGCAGCAGGTTT 1140

Db 1081 CGCAAAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGCGAGCTGGCAGCAGGTTT 1140

Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200

Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200

Qy 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAAGTGGTGGCGGCGGCTTCTGGCG 1260

Db 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAAGTGGTGGCGGCGGCTTCTGGCG 1260

Qy 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGTTCGTAATCACTGCGAATTCG 1320

Db 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGCAGTTCGTAATCACTGCGAATTCG 1320

Qy 1321 TGTGCGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGCGACATCATACGGTT 1380

Db 1321 TGTGCGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGCGACATCATACGGTT 1380

Qy 1381 CTGGCAATATTTCTCAATAGCTCTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440

Db 1381 CTGGCAATATTTCTCAATAGCTCTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440

Qy 1441 ATTGTGAGCGGATAACAATTTTACACAGGAAACAGCCAGTCCGTTTAGTGTTTTTTCAGCA 1500

Db 1441 ATTGTGAGCGGATAACAATTTTACACAGGAAACAGCCAGTCCGTTTAGTGTTTTTTCAGCA 1500

Qy 1501 GCATTTCACCAAGGACCATAGATTATGAAAACTGAAGAAAGGTAACCTGTTAATCTGG 1560

Db 1501 GCATTTCACCAAGGACCATAGATTATGAAAACTGAAGAAAGGTAACCTGTTAATCTGG 1560

Qy 1561 ATTAACGGCGATAAAGGCTATAACGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620

Db 1561 ATTAACGGCGATAAAGGCTATAACGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620

Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT 1680

Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT 1680

Qy 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGACACGACCGTTTGGTGGCTAC 1740

Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGACACGACCGTTTGGTGGCTAC 1740

Qy 1741 GCTCAATCTGGCGCTTGTGGCTGAATCAACCGGACAAAGCGTTCCAGGACAGCTGTAT 1800

Db 1741 GCTCAATCTGGCGCTTGTGGCTGAATCAACCGGACAAAGCGTTCCAGGACAGCTGTAT 1800

Qy 1801 CCGTTTACCTGGGATGCCGTAGCTTACACGCGCAAGCTGATTGCTTACCCGATCGCTGTT 1860

Db 1801 CCGTTTACCTGGGATGCCGTAGCTTACACGCGCAAGCTGATTGCTTACCCGATCGCTGTT 1860

Qy 1861 GAAGCGTTATCGCTGATTTTATACCAAGATCTGCTGCGGAAACCGCCCAAAACCTGGGAA 1920

Db 1861 GAAGCGTTATCGCTGATTTTATACCAAGATCTGCTGCGGAAACCGCCCAAAACCTGGGAA 1920

Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980

Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980

Qy 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGANTGCTGTGACGGGGTTATGCGTTCAAG 2040

Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGANTGCTGTGACGGGGTTATGCGTTCAAG 2040

Qy 2041 TATGAAACCGCAAGTACGACATTAAGAGACGCTGGCGTGGATAACGCTGGCGGAAAGCG 2100

Db 2041 TATGAAACCGCAAGTACGACATTAAGAGACGCTGGCGTGGATAACGCTGGCGGAAAGCG 2100

Qy 2101 GGTCTGACCTTCTCTGTTGACCTGATTTAAAAAACAACAATGACGACACCGATTAC 2160

Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTTAAAAAACAACAATGACGACACCGATTAC 2160

Qy 2161 TCCATCGCAAGCTGCTTCTTAAATAAAGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220

Db 2161 TCCATCGCAAGCTGCTTCTTAAATAAAGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220

Qy 2221 GCATGGTCCAACTCGACACCAAGTGAATTAATGGTGTAAACGGTACTCGGACCTTC 2280

Db 2221 GCATGGTCCAACTCGACACCAAGTGAATTAATGGTGTAAACGGTACTCGGACCTTC 2280

Qy 2281 AAGGTCACACCATCCAAACCGTTCGTTGGCGTGTGAGCGCAGGTATTAACGCGCCAGT 2340

Db 2281 AAGGTCACACCATCCAAACCGTTCGTTGGCGTGTGAGCGCAGGTATTAACGCGCCAGT 2340

Qy 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400

Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400

Qy 2401 GAAGCGGTTAATAAGACAAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460

Db 2401 GAAGCGGTTAATAAGACAAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460

Qy 2461 TTGCGGAAAGTCCACGCTATTGCGGCCACCACTGAAAAACGCCAGAAAAGTGAATCATG 2520

Db 2461 TTGCGGAAAGTCCACGCTATTGCGGCCACCACTGAAAAACGCCAGAAAAGTGAATCATG 2520

Qy 2521 CCGAACATCCCGCAGATGTCGCGCTTTCTGGTATGCGCGTCTGCGGTGATCAACGCC 2580

Db 2521 CCGAACATCCCGCAGATGTCGCGCTTTCTGGTATGCGCGTCTGCGGTGATCAACGCC 2580

Qy 2581 GCCAGCGTCTGACAGCTCGATGAAAGCGCTGAAAGACGCCAGACTAAATTCGAGCTCG 2640

Db 2581 GCCAGCGTCTGACAGCTCGATGAAAGCGCTGAAAGACGCCAGACTAAATTCGAGCTCG 2640

RESULT 15

AF031088 6706 bp DNA circular SYN 07-MAY-2001

LOCUS Shuttle vector pMAL-pIII, complete sequence.

DEFINITION AF031088

ACCESSION AF031088

VERSION AF031088.1 GI:2623821

KEYWORDS Shuttle vector pMAL-pIII

SOURCE Shuttle vector pMAL-pIII

ORGANISM Shuttle vector pMAL-pIII

OTHER SEQUENCES; artificial sequences; vectors.

REFERENCE 1 (bases 1 to 6706)

AUTHORS Zwick,M.B., Bonnycastle,L.L., Noren,K.A., Venturini,S., Leong,E., Barbas,C.F. III, Noren,C.J. and Scott,J.K.

TITLE The maltose-binding protein as a scaffold for monovalent display of peptides derived from phage libraries

JOURNAL Anal. Biochem. 264 (1), 87-97 (1998)

MEDLINE 99002881

PUBMED 9784192

REFERENCE 2 (bases 1 to 6706)

AUTHORS Noren,K.A. and Noren,C.J.

TITLE Direct Submission

JOURNAL Submitted (22-OCT-1997) New England Biolabs, 32 Tozer Road, Beverly, MA 01915, USA

FEATURES

source

1..6706

/organism="Shuttle vector pMAL-pIII"

/mol_type="genomic DNA"

```

/db xref="taxon:69327"
/notes="derived from commercially available vector pMal-p2
(New England Biolabs)"
15..1163
/genes="lacI"
/notes="E. coli gene"
15..46
/genes="lacI"
/notes="lacI"
81..1163
/genes="lacI"
/notes="transcriptional repressor"
/codon_start=1
/transl_table=11
/product="lac repressor"
/protein_id="AAB86558.1"
/db_xref="GI:2623822"
/translation="MKPVTLYDAEYAGYSYQTSRVNQASHVSAKTRKVEAMAE
LNYIPNRVAQLAGKQSLIGVATSSLAHLHAPSQIVAAIKSRADQLGASVVSVMERS
GVBEACAAVHNLDAQRVSGLIINYPLDDDAIAVEARCTNPALFLDYSDQTPINSII
FSDGDTRLGVHVALGHQIALLAGPLSSVSARLRAGLGHKYLTRNQIQPIAREG
DWSAMSGFOOTQMLNEGIVPTAMLVANDQMALGAMRAITBSGLRVGADISVVGVDIT
EDSSCVIPPLTIIKQDFRLIGQTSVDRLLQLSQGAVKGNQLLPVSLVYRKTTIAPNT
QTASPRALADSLMLQARQVSRLESQ"
1406..1433
/notes="Pac promoter"
1528..2976
/genes="male-lacZa fusion"
/notes="E. coli gene fusion"
1528..2976
/genes="malB-lacZa fusion"
/notes="MBP-lacZa; MBP modified to include phage M13
secretion signal sequence"
/codon_start=1
/transl_table=11
/product="maltose binding protein-lacZ alpha peptide
fusion protein precursor"
/protein_id="AAB86559.1"
/db_xref="GI:2623823"
/translation="MKKLLFAIPLVVPFYSHADMKIEEGKLIWIINDKGYGLAEV
GKFEKDTGIVTVVHEDPKLEKFPQVAATGDDPIIFWADHDFGQVAGSGLLAEIIP
DKAFQDKLYPTDAVRNKGIIAYIAVEALSLIYNKDLNPNPKTWEEI PALDKEL
KAKGSLAMFNLQBPYFWPLIAADGGVAPKVKYKVDIKDVGVDNAGAKAGLFLVD
LKNKNMADTYSIAEAPNKGEMTAMTINGPMWSNIDTSKNYGVTVLPFKQPS
KPFVGLSAGINASPKNELAKELENYLLTDEGLEAVNKDKPLGVALKSYBEELAK
DPRVLAATHNAQKEINPFIQMSAFWAVTAIVNAASGRQTVDEALRDAQTNSSN
NNNNNNLIGRISEFSGSRVDLQASLALAVLQLRDWNENPGVTQLNRLAAHPFP
ASWRNSEARTDRPSQQLRSLSNGEWQLGCFG"
1528..1581
/genes="malE-lacZa fusion"
/notes="M13 secretion signal sequence"
1582..2973
/genes="malB-lacZa fusion"
/product="maltose binding protein-lacZ alpha peptide
fusion protein"
3478..4338
/genes="bla"
/notes="R-plasmid ampicillin resistance gene"
3478..4338
/genes="bla"
/notes="penicillinase"
/codon_start=1
/transl_table=11
/product="RTEM-1 beta-lactamase"
/protein_id="AAB86560.1"
/db_xref="GI:2623824"
/translation="MSIQHPRVRLIIPFAAFCLPVFAHPETLVKVKDAEDQLGARVGY
IEDLSNGKILSEFRPEERPMSTFKVLLCGAVLSRVDAQOQLGRLRIHYSQNDLVE
YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTIGPKELTAFLHNMGDHVTPL
DRWPELNEAIPIQDERDTTPVAMATTLKRLTGLLTSLASRQLLDHWEADKIVAGL
LRGALPAGWFIADKSGAGERSGRIIAALGPDGPKSRIVVIYITGTSQATWDERNROIA
EIGASLIKHW"
4381..4893
/standard_name="M13 origin (+ -)"

```

```

ORIGIN
Query Match
Best Local Similarity 79.7%; Score 2631; DB 12; Length 6706;
Matches 2704; Conservative 0; Mismatches 5; Indels 60; Gaps 1;

QY 1 CGACACCATCGAATGGTGCAGAAACCTTTTCGCGTATGGCATGATAGCGCCCGGAAGAGA 60
DB 1 CGACACCATCGAATGGTGCAGAAACCTTTTCGCGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCAAATTCAGGCGTGTGAATGTGAACCAAGTAAAGTTATACGATGTGCGAGAGTATGCCG 120
DB 61 GTCAAATTCAGGCGTGTGAATGTGAACCAAGTAAAGTTATACGATGTGCGAGAGTATGCCG 120
QY 121 GTGCTCTTTATCAGACCGCTTTTCCCGCGTGTGAACCAAGCGCAGCACGTTTCTCGCAAAA 180
DB 121 GTGCTCTTTATCAGACCGCTTTTCCCGCGTGTGAACCAAGCGCAGCACGTTTCTCGCAAAA 180
QY 181 CGCGGGAAAAGTGAAGCGCGCATGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
DB 181 CGCGGGAAAAGTGAAGCGCGCATGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGCGCGTTGCCACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGCGCGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTGCAGAAATTTGTCGCGCGATTAAATCTCGCGCGATCAACTGGGTGCCAGCG 360
DB 301 ACGCGCGTGCAGAAATTTGTCGCGCGATTAAATCTCGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTGTGATGTAGAACGAGCGGTGCAAGCGCTGTAAGCGCGGTGCAACAATC 420
DB 361 TGGTGGTGTGATGTAGAACGAGCGGTGCAAGCGCTGTAAGCGCGGTGCAACAATC 420
QY 421 TTCTCGCGCAACCGTCAGTGGGCTGATCATTAATCATCTCGCTGGATGACACAGATGCCA 480
DB 421 TTCTCGCGCAACCGTCAGTGGGCTGATCATTAATCATCTCGCTGGATGACACAGATGCCA 480
QY 481 TTGCTGTGGAGCTGCTGCACTAAATGTTCCGCGCTTATTTCTTGATGTCTCTGACCA 540
DB 481 TTGCTGTGGAGCTGCTGCACTAAATGTTCCGCGCTTATTTCTTGATGTCTCTGACCA 540
QY 541 CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGGACTGGCGGTGGAGCATC 600
DB 541 CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGGACTGGCGGTGGAGCATC 600
QY 601 TGTTCGATTTGGTGCACGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
DB 601 TGTTCGATTTGGTGCACGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTGCTGCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCGCATAG 720
DB 661 CGCGTCTGCTGCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCGCATAG 720
QY 721 CGGAACGGGAAGCGCACTGGAGTGCCATGTCGCGTTTTTCAACAAACCATGCAAAATGCTGA 780
DB 721 CGGAACGGGAAGCGCACTGGAGTGCCATGTCGCGTTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTCCCACTCGCATGCTGGTTGCCAAGATCAGATGGCGTGGCGCA 840
DB 781 ATGAGGGCATCGTCCCACTCGCATGCTGGTTGCCAAGATCAGATGGCGTGGCGCA 840
QY 841 TCGCGGCCATTACCGAGTCCGGCTGGCGTTGGTGGCATATCTCGGTAGTGGGATAG 900
DB 841 TCGCGGCCATTACCGAGTCCGGCTGGCGTTGGTGGCATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
QY 961 GCCTCTCGGGCAAAACCAAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
DB 961 GCCTCTCGGGCAAAACCAAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020

```

Qy	1021	AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAAAACCAACCTCGCGGCCAATA	1080
Db	1021	AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAAAACCAACCTCGCGGCCAATA	1080
Qy	1081	CGCAAAACCGCTCTCCCGCGGTTGGCCGATTCAATTAATGAGCTGGCAGACAGAGTTT	1140
Db	1081	CGCAAAACCGCTCTCCCGCGGTTGGCCGATTCAATTAATGAGCTGGCAGACAGAGTTT	1140
Qy	1141	CCGACATGGAAGCGGCGAGTGAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTAG	1200
Db	1141	CCGACATGGAAGCGGCGAGTGAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTAG	1200
Qy	1201	GCACAAATCTCATGTTTGACAGCTTATCATGACGCGGTGACACCAATGCTTCTGCGC	1260
Db	1201	GCACAAATCTCATGTTTGACAGCTTATCATGACGCGGTGACACCAATGCTTCTGCGC	1260
Qy	1261	TCAGGACGCCATCGGAAGCTGTGGTATGCGTGTGCGAGTCTGTAATCACTGCATAATTTCG	1320
Db	1261	TCAGGACGCCATCGGAAGCTGTGGTATGCGTGTGCGAGTCTGTAATCACTGCATAATTTCG	1320
Qy	1321	TGTCGCTCAAGCGCACCTCCCGTCTCGGATTAATGTTTTTGGCGGACATCATACGGTT	1380
Db	1321	TGTCGCTCAAGCGCACCTCCCGTCTCGGATTAATGTTTTTGGCGGACATCATACGGTT	1380
Qy	1381	CTGGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGTTGA	1440
Db	1381	CTGGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGTTGA	1440
Qy	1441	ATTGTGAGCGGATAACAAATTTACACAGGAACAGCCAGTCCGTTTAGTGTGTTTACAGA	1500
Db	1441	ATTGTGAGCGGATAACAAATTTACACAGGAACAGCCAGTCCGTTTAGTGTGTTTACAGA	1500
Qy	1501	GCACCTTACCACAGGACCATAG-----	1524
Db	1501	GCACCTTACCACAGGACCATAGCATATGAAAAATTTATTTCGCAATTCCTTTAGTG	1560
Qy	1525	-----ATTATGAAAACTGAAGAGGTAAACTGGTAATCTGG	1560
Db	1561	GTAACCTTTCTCACTCGGCCGATATGAAAAATCGAAGAGGTAAACTGGTAATCTGG	1620
Qy	1561	ATTTAACGGGATTAAGGCTTAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT	1620
Db	1621	ATTTAACGGGATAAAGGCTTAACCGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT	1680
Qy	1621	ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACACAGGT	1680
Db	1681	ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACACAGGT	1740
Qy	1681	CGGCAATGCGGATGGCCCTGACATTTATCTTGGGCACACGACCGCTTTGGTGGCTTAC	1740
Db	1741	CGGCAATGCGGATGGCCCTGACATTTATCTTGGGCACACGACCGCTTTGGTGGCTTAC	1800
Qy	1741	GCTCAATCTGGCCTGTTGGCTGAATATCAACCGGCAAGAGTTCAGGACAGCTGTAT	1800
Db	1801	GCTCAATCTGGCCTGTTGGCTGAATATCAACCGGCAAGAGTTCAGGACAGCTGTAT	1860
Qy	1801	CGTTTACCTGGGATGCGGTACGTTTACAAAGCGCAAGCTGATTGCTTACCGGATCGCTGT	1860
Db	1861	CGTTTACCTGGGATGCGGTACGTTTACAAAGCGCAAGCTGATTGCTTACCGGATCGCTGT	1920
Qy	1861	GAAGCGTTATCGCTGATTTATTAACAAAGATCTGCTGCCGAACCCCGCAAAACCTGGGAA	1920
Db	1921	GAAGCGTTATCGCTGATTTATTAACAAAGATCTGCTGCCGAACCCCGCAAAACCTGGGAA	1980
Qy	1921	GAGATCCCGGCTCGGATAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	1980
Db	1981	GAGATCCCGGCTCGGATAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC	2040
Qy	1981	CTGCAAGAACCGTACTTACCTGGCGGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG	2040
Db	2041	CTGCAAGAACCGTACTTACCTGGCGGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG	2100

Qy	2041	TATGAAAACGGCAAGTACGACATTTAAAGAGCTGGCGTGGATACGCTGGCGCGAAGCG	2100
Db	2101	TATGAAAACGGCAAGTACGACATTTAAAGAGCTGGCGTGGATACGCTGGCGCGAAGCG	2160
Qy	2101	GGTCTGACCTTCCCTGGTTGACCTGATTTAAAAACAAACACATGAATGCAGACACCGATTAC	2160
Db	2161	GGTCTGACCTTCCCTGGTTGACCTGATTTAAAAACAAACACATGAATGCAGACACCGATTAC	2220
Qy	2161	TCCATCGAGAGAGCTGCTCTTTTAAAGGCGAAACAGCGATGACATCAACAGCGCCGTGG	2220
Db	2221	TCCATCGAGAGAGCTGCTCTTTTAAAGGCGAAACAGCGATGACATCAACAGCGCCGTGG	2280
Qy	2221	GCATGGTCCACATCGACACAGCAAGTGAATTTATGGTTAAACGGTACTCGCGACCTTC	2280
Db	2281	GCATGGTCCACATCGACACAGCAAGTGAATTTATGGTTAAACGGTACTCGCGACCTTC	2340
Qy	2281	AAGGTCACACCATCCAAACCGTTTGGTGGCTGTGAGCGCAGGTATTTAAACCGCCAGT	2340
Db	2341	AAGGTCACACCATCCAAACCGTTTGGTGGCTGTGAGCGCAGGTATTTAAACCGCCAGT	2400
Qy	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGTCTG	2400
Db	2401	CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGTCTG	2460
Qy	2401	GAAGCGGTTAATAAGACAAACCGCTGGTGGCTAGCGCTGAAGTCTTACGAGGAAG	2460
Db	2461	GAAGCGGTTAATAAGACAAACCGCTGGTGGCTAGCGCTGAAGTCTTACGAGGAAG	2520
Qy	2461	TTGGCGAAAGATCCACGCTATTTCGCCACCATGGAACACGCCAGAAAGGTGAATCATG	2520
Db	2521	TTGGCGAAAGATCCACGCTATTTCGCCACCATGGAACACGCCAGAAAGGTGAATCATG	2580
Qy	2521	CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAACGCC	2580
Db	2581	CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAACGCC	2640
Qy	2581	GCCAGCGTCTGACAGCTGTGATGAAGCCCTGAAAGCGCGCAGACTAATTCGAGCTCG	2640
Db	2641	GCCAGCGTCTGACAGCTGTGATGAAGCCCTGAAAGCGCGCAGACTAATTCGAGCTCG	2700
Qy	2641	AACAAACAAACAATAACAATAACAACAACTCGGGATCGAGGGAAGGATTTCAAGATTC	2700
Db	2701	AACAAACAAACAATAACAATAACAACAACTCGGGATCGAGGGAAGGATTTCAAGATTC	2760
Qy	2701	GGATCCTCT 2709	
Db	2761	GGATCCTCT 2769	

Search completed: August 1, 2005, 17:22:13
Job time : 9284 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 14:24:04 ; Search time 1094.6 Seconds
(without alignment)
17846.836 Million cell updates/sec

Title: US-09-765-555B-18
Perfect score: 3300
Sequence: 1 ccgacaccatgaatggtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3300	100.0	3300	4 AAD11592	Aad11592 Partial s
2	3232.8	98.0	3300	4 AAD11591	Aad11591 Partial s
3	3226.4	97.8	3300	4 AAD11588	Aad11588 Partial s
4	3226.4	97.8	3300	4 AAD11589	Aad11589 Partial s
5	3216.8	97.5	3300	4 AAD11590	Aad11590 Partial s
6	2705.8	82.0	8101	4 AAH25681	AAH25681 Nucleotid
7	2705.8	82.0	8101	6 ABL53238	ABL53238 Nucleotid
8	2690.4	81.5	7259	12 ADO23608	ADO23608 DNA encod
9	2690.4	81.5	7322	12 ADO23613	ADO23613 DNA encod
10	2690.4	81.5	7352	12 ADO23603	ADO23603 DNA encod
11	2690.4	81.5	7370	12 ADO23598	ADO23598 DNA encod
12	2690.4	81.5	7370	12 ADO23639	ADO23639 DNA encod
13	2690.4	81.5	7370	12 ADO23649	ADO23649 DNA encod
14	2690.4	81.5	7370	12 ADO23644	ADO23644 DNA encod
15	2690.4	81.5	7403	12 ADO23594	ADO23594 DNA encod
16	2690.4	81.5	7442	12 ADO23590	ADO23590 DNA encod
17	2690.4	81.5	7478	12 ADO23584	ADO23584 DNA encod
18	2688.8	81.5	7112	12 ADO23618	ADO23618 DNA encod
19	2636.8	79.9	6806	6 ABL49925	ABL49925 Maltose b
20	2605.4	79.0	7553	12 ADO23588	ADO23588 DNA encod

21	1609.4	48.8	5558	12	ADL90419	Adl90419 Clostridi
22	1474	44.7	4700	9	ACF06053	Act06053 Vector pJ
23	1472.4	44.6	4700	12	ADL18670	Adl18670 Vector pJ
24	1462.4	44.3	4920	10	ADH73599	Adh73599 Plasmid p
25	1462.4	44.3	4935	10	ADH73598	Adh73598 Plasmid p
26	1462.4	44.3	4945	10	ADH73600	Adh73600 Plasmid p
27	1462.4	44.3	4951	10	ADH73601	Adh73601 Plasmid p
28	1418.2	43.0	5903	8	ABZ23939	Abz23939 Nucleotid
29	1417.6	43.0	1511	6	ABQ73200	Abq73200 E. coli t
30	1417.6	43.0	5926	2	AAV32977	Aav32977 Tn7 donor
31	1417.6	43.0	5926	6	AAV32977	Aav32977 Tn7 donor
32	1417.6	43.0	5926	6	AAV32977	Aav32977 Tn7 donor
33	1362	41.3	5201	12	ADL72229	Adl72229 DNA seque
34	1362	41.3	5201	12	ADL72228	Adl72228 DNA seque
35	1238.4	37.5	5024	9	ACF06051	Act06051 Plasmid p
36	1238.4	37.5	5024	12	ADL18668	Adl18668 Plasmid p
37	1237.2	37.5	1922	6	ABQ73089	Abq73089 MBP and z
38	1237.2	37.5	1922	10	ADF83622	Adf83622 MBP-human
39	1204	36.5	4557	2	AAH90491	Aah90491 Vector pl
40	1204	36.5	4969	12	ADO33522	Ado33522 Expressio
41	1204	36.5	5006	10	ADE24112	Ade24112 plasmid p
42	1204	36.5	5064	12	ADO33532	Ado33532 Vector pG
43	1204	36.5	5082	12	ADO33531	Ado33531 Vector pG
44	1204	36.5	6259	6	AAO29720	Aao29720 Plasmid p
45	1204	36.5	6823	3	AAC55459	Aac55459 Destinati

ALIGNMENTS

RESULT 1
AAD11592
ID AAD11592 standard; DNA; 3300 BP.
XX
AC AAD11592;
XX
DT 24-SEP-2001 (first entry)
XX
DE Partial sequence of pMal-Ap3 and ZFPAP3 DNA.
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
KW modulation; plant technology; agriculture; Ap3; APETALA3; ds.
XX
OS Unidentified.
XX
FH Key
CDS Location/Qualifiers
FT 2719..3270
FT /*tag= a
FT /product= "ZFPAP3 protein"
FT /transl_except= {pos:3124..3126, aa:Ser)
FT /transl_except= {pos:3127..3129, aa:Ser)
FT /transl_except= {pos:3136..3138, aa:Val)
FT /note= "CDS does not include start and stop codon"
FT /partial
FT primer_bind 2740..2790
FT /*tag= b
FT /bound_moiety= "F1-f2 primer"
FT primer_bind 2770..2850
FT /*tag= c
FT /bound_moiety= "F1-f1 primer"
FT primer_bind 2824..2889
FT /*tag= d
FT /bound_moiety= "F2-b primer"
FT primer_bind 2867..2940
FT /*tag= e
FT /bound_moiety= "F2-f primer"
FT primer_bind 2916..2973
FT /*tag= f
FT /bound_moiety= "F3-b1 primer"
FT primer_bind 2953..3021
FT /*tag= g
FT /bound_moiety= "F3-b2 primer"
FT primer_bind 2992..3042


```
FT      /*tag= h
FT      /bound_moiety= "F4-f2 primer"
FT      3022..3102
FT      /*tag= i
FT      /bound_moiety= "F4-f1 primer"
FT      3076..3141
FT      /*tag= j
FT      /bound_moiety= "F5-b primer"
FT      3119..3192
FT      /*tag= k
FT      /bound_moiety= "F5-f primer"
FT      3168..3225
FT      /*tag= l
FT      /bound_moiety= "F6-b1 primer"
FT      3205..3273
FT      /*tag= m
FT      /bound_moiety= "F6-b2 primer"
XX
PN      WO200152620-A2.
PD      26-JUL-2001.
XX
XX      19-JAN-2001; 2001WO-US001817.
PF
PR      21-JAN-2000; 2000US-0177468P.
PR      21-JUL-2000; 2000US-00620897.
XX
PA      (SCRI ) SCRIPPS RES INST.
PA      (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
PI      Barbas CF, Stege JT, Guan X, Dalmia B;
XX      WPI; 2001-465325/50.
DR      P-PSDB; AA06005.
XX
XX      New zinc finger proteins, useful for modulating or regulating gene
PT      expression and metabolic pathways in plants, e.g. for treating in the
PT      plant cells a disorder that is associated with abnormal expression of the
PT      target gene.
PS      Example 4; Page 148-149; 156pp; English.
XX
XX      The patent discloses methods and compositions to modulate the expression
CC      of a target gene in plant cells. The method involves providing plant
CC      cells with a zinc finger protein (ZFP) which is capable of specifically
CC      binding to a target nucleotide sequence or its complementary strand
CC      within a target gene and allowing the ZFP binding to the target
CC      nucleotide sequence, where the expression of the target gene in the plant
CC      cells is modulated. The ZFP and fusions of the ZFP proteins are useful
CC      for modulating or regulating gene expression and metabolic pathways in
CC      plants. The ZFP, fusion proteins and methods are useful in plant and
CC      agricultural technology. The method is useful particularly for treating a
CC      disorder in the plant cells, where the disorder is associated with
CC      abnormal expression of the target gene. The present DNA sequence is the
CC      partial sequence of pMal-AP3 (AP3FLA3) and ZFPp3 DNA
XX
SQ      Sequence 3300 BP; 853 A; 872 C; 876 G; 699 T; 0 U; 0 Other;
Query Match
Best Local Similarity 100.0%; Score 3300; DB 4; Length 3300;
Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 CCGACACCATCGAATGGTGC AAAACCTTTCGCGTATGGCATGATAGCCCGCGGAAGAGA 60
Db      1 CCGACACCATCGAATGGTGC AAAACCTTTCGCGTATGGCATGATAGCCCGCGGAAGAGA 60
Qy      61 GTCAATTCAGGTGGTGAATGTGAACACGATTAACGTTATACGATGTCGAGATATCCG 120
Db      61 GTCAATTCAGGTGGTGAATGTGAACACGATTAACGTTATACGATGTCGAGATATCCG 120
Qy      121 GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCAGTTCCTGCGAAAA 180
Db      121 GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCAGTTCCTGCGAAAA 180
```

```
Qy      181 CGCGGGAAAAGTCGAAGCGCGGATGGGAGCTGAATTACATTCCCAACCGCGTGGCAC 240
Db      181 CGCGGGAAAAGTCGAAGCGCGGATGGGAGCTGAATTACATTCCCAACCGCGTGGCAC 240
Qy      241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGCGCCCTGC 300
Db      241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGCGCCCTGC 300
Qy      301 ACGCGCGTGC AAAATTGTCCGCGGATTAATCTCCGCGCGATCAACTGGGTGCCAGCG 360
Db      301 ACGCGCGTGC AAAATTGTCCGCGGATTAATCTCCGCGCGATCAACTGGGTGCCAGCG 360
Qy      361 TGGTGGTGTGATAGCGAGCGCGCTGGAAGCGCTGTAAGCGCGCGTGCACAATC 420
Db      361 TGGTGGTGTGATAGCGAGCGCGCTGGAAGCGCTGTAAGCGCGCGTGCACAATC 420
Qy      421 TTCTCGCGCAACCGCTCAGTGGGCTGATCACTTAACCTATCCGCTGATGACCCAGATGCCA 480
Db      421 TTCTCGCGCAACCGCTCAGTGGGCTGATCACTTAACCTATCCGCTGATGACCCAGATGCCA 480
Qy      481 TTGCTGTGGAAGCTGCGCTGCACTAAATGTTCCGGCGTTAATTTCTTGATGTCTCTGACCAGA 540
Db      481 TTGCTGTGGAAGCTGCGCTGCACTAAATGTTCCGGCGTTAATTTCTTGATGTCTCTGACCAGA 540
Qy      541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Db      541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Qy      601 TGGTTCGATTGGGTACCAGCAAAATCGCGCTGTTAGCGGGCCCCATTAAAGTTCTGTCTCGG 660
Db      601 TGGTTCGATTGGGTACCAGCAAAATCGCGCTGTTAGCGGGCCCCATTAAAGTTCTGTCTCGG 660
Qy      661 CGGCTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db      661 CGGCTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy      721 CGGAACGGGAAGCGCATGCGAGTCCCATGTCGGTTCCTCAACCAACCATCAAAATGCTGA 780
Db      721 CGGAACGGGAAGCGCATGCGAGTCCCATGTCGGTTCCTCAACCAACCATCAAAATGCTGA 780
Qy      781 ATGAGGCAATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
Db      781 ATGAGGCAATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
Qy      841 TGGCGCCATTACCGAGTCCGGCTCGGCTGGTGGCGGATATCTCGGTAGTGGGATACG 900
Db      841 TGGCGCCATTACCGAGTCCGGCTCGGCTGGTGGCGGATATCTCGGTAGTGGGATACG 900
Qy      901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAACCCACCATCAAAACAGATTTTC 960
Db      901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAACCCACCATCAAAACAGATTTTC 960
Qy      961 GCCTGTGGGGCAAAACAGCGTGGACCGTGTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db      961 GCCTGTGGGGCAAAACAGCGTGGACCGTGTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Qy      1021 AGGGCAATCAGCTGTGGCGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCCAATA 1080
Db      1021 AGGGCAATCAGCTGTGGCGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCCAATA 1080
Qy      1081 CGCAAAACCGCTCTCCCGCGCGTGGCGGATTCATTAAATGCAGCTGGGACGACAGGTTT 1140
Db      1081 CGCAAAACCGCTCTCCCGCGCGTGGCGGATTCATTAAATGCAGCTGGGACGACAGGTTT 1140
Qy      1141 CCCGACTGGAAGCGGGCGAGTGAAGCGCAACGCAATTAATGTGAGTTAGTCACTCAATTAG 1200
Db      1141 CCCGACTGGAAGCGGGCGAGTGAAGCGCAACGCAATTAATGTGAGTTAGTCACTCAATTAG 1200
Qy      1201 GCACAATTCCTATGTTTGAAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260
Db      1201 GCACAATTCCTATGTTTGAAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260
```

```
QY 1261 TCAGGACGACATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTAAATCACTGCATTAATTGG 1320
Db 1261 TCAGGACGACATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTAAATCACTGCATTAATTGG 1320
QY 1321 TGTCCCTCAAGCGGCACATCCCGCTTCCTGGATATGTTTTTGGCCGACATCATAAACGGTT 1380
Db 1321 TGTCCCTCAAGCGGCACATCCCGCTTCCTGGATATGTTTTTGGCCGACATCATAAACGGTT 1380
QY 1381 CTGGCAATATTTCTGAATAGCTGTTGACAAATTAATCATTCGGCTCGTATAATGTTGTGA 1440
Db 1381 CTGGCAATATTTCTGAATAGCTGTTGACAAATTAATCATTCGGCTCGTATAATGTTGTGA 1440
QY 1441 ATTGTAGCGGATACAAATTTTACACAGAAACACCCAGTCCGTTAGTGTTTTTTCAGA 1500
Db 1441 ATTGTAGCGGATACAAATTTTACACAGAAACACCCAGTCCGTTAGTGTTTTTTCAGA 1500
QY 1501 GCACCTTCCACCAAGGACCATAGATTATGAAACTGGAAGAGTAAACTGGTAATCTGG 1560
Db 1501 GCACCTTCCACCAAGGACCATAGATTATGAAACTGGAAGAGTAAACTGGTAATCTGG 1560
QY 1561 ATTAAACGGGATAAAGGCTATAACCGTCTCGTGAAGTCGTTAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAAACGGGATAAAGGCTATAACCGTCTCGTGAAGTCGTTAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGATAAACTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGATAAACTGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACACCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTTGGCTGAATATCAACCCCGGACAAAGCGTTCCAGACAAAGCTGAT 1800
Db 1741 GCTCAATCTGGCTGTTGGCTGAATATCAACCCCGGACAAAGCGTTCCAGACAAAGCTGAT 1800
QY 1801 CCGTTTACCTGGATGCGGTAGCTTACACGCAAGCTGATTGCTTACCCGATCCCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGGTAGCTTACACGCAAGCTGATTGCTTACCCGATCCCTGTT 1860
QY 1861 GAAGGTTATCGCTGATTTATAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGGTTATCGCTGATTTATAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCTCGATAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCTCGATAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGGCTGATTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGGCTGATTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACGGCAAGTACGACATTAAGACGTGGGCGTGATTAACGTCGGCGGAAAGCG 2100
Db 2041 TATGAAACGGCAAGTACGACATTAAGACGTGGGCGTGATTAACGTCGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAAGAAACAAACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAAGAAACAAACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAAGAGTCCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAAGAGTCCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCACATCGACACGCAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCACATCGACACGCAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCAAAACCGTTCTGTTGGCGTGTGAGCGCAGGATTAACGCCCGCAGT 2340
Db 2281 AAGGGTCAACCATCAAAACCGTTCTGTTGGCGTGTGAGCGCAGGATTAACGCCCGCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAAACTATCTGCTGACTGATGAAGGTCGT 2400
```

```
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAAACTATCTGCTGACTGATGAAGGTCGT 2400
QY 2401 GAAGCGGTTAATAAGACAAACCCGCTGGGTCCGCTAGCGCTGAAGTCTTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCCGCTGGGTCCGCTAGCGCTGAAGTCTTTACGAGGAAG 2460
QY 2461 TTGGGAAAGATCCACGTTATTTGGCCGACCAATGGAANAACCCGAGAAAGGTAATCATG 2520
Db 2461 TTGGGAAAGATCCACGTTATTTGGCCGACCAATGGAANAACCCGAGAAAGGTAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCTGCTGCTGCTGCTGCTGCTG 2580
QY 2581 GCCAGCGTCTGCTGAGATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTG 2640
Db 2581 GCCAGCGTCTGCTGAGATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTG 2640
QY 2641 AACAAACAAACAAATAACAATAACAACCTCGGGATCGAGGGAAGGATTTCAAAATTC 2700
Db 2641 AACAAACAAACAAATAACAATAACAACCTCGGGATCGAGGGAAGGATTTCAAAATTC 2700
QY 2701 GGATCCTCTCTCTGCTGGCCGACGCGCCCTCGAGCCCGGGGAGAGCCCTATGCTGT 2760
Db 2701 GGATCCTCTCTCTGCTGGCCGACGCGCCCTCGAGCCCGGGGAGAGCCCTATGCTGT 2760
QY 2761 CCGGAATCTGCTGTAAGTCTCTCAGCCAGACGAGTCCCTGCTGCTGCGCCACGAGCTACCCAC 2820
Db 2761 CCGGAATCTGCTGTAAGTCTCTCAGCCAGACGAGTCCCTGCTGCTGCGCCACGAGCTACCCAC 2820
QY 2821 ACGGTTGAAAAACCGGTATAAATGCCAGAGTGGGCAAACTCTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGTTGAAAAACCGGTATAAATGCCAGAGTGGGCAAACTCTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGTCGCGCCATCAACGACATCATCTGCGGAGAGCCATACAAATGTCAGAAATGTGGC 2940
Db 2881 CTGTCGCGCCATCAACGACATCATCTGCGGAGAGCCATACAAATGTCAGAAATGTGGC 2940
QY 2941 AAGTCTTTTTCAGCCAGTCCAGCAACCTGCTGGGCCACCAACGTAATCTCACACCGGGGAGAG 3000
Db 2941 AAGTCTTTTTCAGCCAGTCCAGCAACCTGCTGGGCCACCAACGTAATCTCACACCGGGGAGAG 3000
QY 3001 CCTATGCTTGTCCGGAATGTGTAAGTCTCTTTCAGCAACAGTGGCTCTCTGTTTACACAC 3060
Db 3001 CCTATGCTTGTCCGGAATGTGTAAGTCTCTTTCAGCAACAGTGGCTCTCTGTTTACACAC 3060
QY 3061 CAGGTATCCACACGCGGTGAABACCGGTATAATGCGCCAGAGTGGCGGCAAACTCTTTTACG 3120
Db 3061 CAGGTATCCACACGCGGTGAABACCGGTATAATGCGCCAGAGTGGCGGCAAACTCTTTTACG 3120
QY 3121 CAGCGCGCCACCTGGAAACGCGCATCAACGCGACTCATCTGCGGAGAGCCATACAAATGT 3180
Db 3121 CAGCGCGCCACCTGGAAACGCGCATCAACGCGACTCATCTGCGGAGAGCCATACAAATGT 3180
QY 3181 CCAGATGTGCAAGTCTTTCTCAACTTTTTCAGCAACTTTGGTCCGTCACCAACGTAATCTAC 3240
Db 3181 CCAGATGTGCAAGTCTTTCTCAACTTTTTCAGCAACTTTGGTCCGTCACCAACGTAATCTAC 3240
QY 3241 ACCGTTAAAAAACTAGTGGCCGACGCGCCGACGTAACCGTACGAGTTCCGAGCTACGCT 3300
Db 3241 ACCGTTAAAAAACTAGTGGCCGACGCGCCGACGTAACCGTACGAGTTCCGAGCTACGCT 3300
```

RESULT 2

AAD11591

ID AAD11591 standard; DNA; 3300 BP.

XX AAD11591;

XX AC

DT 24-SEP-2001 (first entry)

XX

DE Partial sequence of pMal-m4 and 2FPm4 DNA.

XX Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
 KW modulation; plant technology; agriculture; ds.
 XX Unidentified.

Key

CDS

Location/Qualifiers
 2719..3270

/*tag= a

/product= "ZFPm4 protein"

/transl_except= (pos:3046..3048, aa:Ser)

/note= "CDS does not include start and stop codon"

/partial

2740..2790

/*tag= b

/bound_moiety= "F1-f2 primer"

2770..2850

/*tag= c

/bound_moiety= "F1-f1 primer"

2824..2889

/*tag= d

/bound_moiety= "F2-b primer"

2867..2940

/*tag= e

/bound_moiety= "F2-f primer"

2916..2973

/*tag= f

/bound_moiety= "F3-b1 primer"

2953..3021

/*tag= g

/bound_moiety= "F3-b2 primer"

2992..3042

/*tag= h

/bound_moiety= "F4-f2 primer"

3022..3102

/*tag= i

/bound_moiety= "F4-f1 primer"

3076..3141

/*tag= j

/bound_moiety= "F5-b primer"

3119..3192

/*tag= k

/bound_moiety= "F5-f primer"

3168..3225

/*tag= l

/bound_moiety= "F6-b1 primer"

3205..3273

/*tag= m

/bound_moiety= "F6-b2 primer"

WO200152620-A2.

26-JUL-2001.

19-JAN-2001; 2001WO-US001817.

21-JAN-2000; 2000US-0177468P.

21-JUL-2000; 2000US-00620897.

(SCRI) SCRIPPS RES INST.

(SYGN) SYNGENTA AGRIC DISCOVERY INC.

Barbas CF, Stege JT, Guan X, Dalmia B;

WPI; 2001-465325/50.

P-PSDB; AAE06004.

New zinc finger proteins, useful for modulating or regulating gene expression and metabolic pathways in plants, e.g. for treating in the plant cells a disorder that is associated with abnormal expression of the target gene.

Example 4; Page 145-147; 156pp; English.

XX The patent discloses methods and compositions to modulate the expression
 CC of a target gene in plant cells. The method involves providing plant
 CC cells with a zinc finger protein (ZFP) which is capable of specifically
 CC binding to a target nucleotide sequence or its complementary strand
 CC within a target gene and allowing the ZFP binding to the target
 CC nucleotide sequence, where the expression of the target gene in the plant
 CC cells is modulated. The ZFP and fusions of the ZFP proteins are useful
 CC for modulating or regulating gene expression and metabolic pathways in
 CC plants. The ZFP, fusion proteins and methods are useful in plant and
 CC agricultural technology. The method is useful particularly for treating a
 CC disorder in the plant cells, where the disorder is associated with
 CC abnormal expression of the target gene. The present DNA sequence is the
 CC partial sequence of pMal-m4 and ZFPm4 DNA
 XX

SQ Sequence 3300 BP; 849 A; 865 C; 880 G; 706 T; 0 U; 0 Other;

Query Match 98.0%; Score 3232.8; DB 4; Length 3300;

Best Local Similarity 98.7%; Pred. No. 0;

Matches 3258; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy	1	CCGACACCATCGAATGGTGC	AAAACCTTTCGCGGTATGG	CATGATAGCGCCCGAAGAGA	60
Db	1	CCGACACCATCGAATGGTGC	AAAACCTTTCGCGGTATGG	CATGATAGCGCCCGAAGAGA	60
Qy	61	GTCAATTACGGTGGTGAATG	TGAACACCTGTAACGTTAT	ACGATGTCGAGATATGCCG	120
Db	61	GTCAATTACGGTGGTGAATG	TGAACACCTGTAACGTTAT	ACGATGTCGAGATATGCCG	120
Qy	121	GTGTCTCTTATCAGACCGT	TTCCCGGTGGTGAACACG	AGCCAGCGTTCCTGCGAAAA	180
Db	121	GTGTCTCTTATCAGACCGT	TTCCCGGTGGTGAACACG	AGCCAGCGTTCCTGCGAAAA	180
Qy	181	CGCGGAAAAGTGGAAAGCG	CGATGCGGAGCTGAATTAC	ATTCCCAACCGCGTGGCAC	240
Db	181	CGCGGAAAAGTGGAAAGCG	CGATGCGGAGCTGAATTAC	ATTCCCAACCGCGTGGCAC	240
Qy	241	AACAACTGGCGGCAAAAC	AGTCTGCTGATTCGGCTG	CCACTTCCCAACCGCGTGG	300
Db	241	AACAACTGGCGGCAAAAC	AGTCTGCTGATTCGGCTG	CCACTTCCCAACCGCGTGG	300
Qy	301	ACGCGCGTGC	CGAAATTTGTCGCGCGAT	TAAATCTCGCGCGATCACT	360
Db	301	ACGCGCGTGC	CGAAATTTGTCGCGCGAT	TAAATCTCGCGCGATCACT	360
Qy	361	TGTTGTTGTCGATGGTAG	ACGAGCGGTGGAAGCGGT	GTAAGCGGTGTAAGCGGTG	420
Db	361	TGTTGTTGTCGATGGTAG	ACGAGCGGTGGAAGCGGT	GTAAGCGGTGTAAGCGGTG	420
Qy	421	TTCTCGCAACCGCTCAGT	GGGCTGATCATTAATCTAT	CGCTGGATGACGAGATGCCA	480
Db	421	TTCTCGCAACCGCTCAGT	GGGCTGATCATTAATCTAT	CGCTGGATGACGAGATGCCA	480
Qy	481	TTGCTGTGGAAGCTGCTGC	ACTAATGTTCCGCGGTATT	CTTGATGTCCTGACAGAGA	540
Db	481	TTGCTGTGGAAGCTGCTGC	ACTAATGTTCCGCGGTATT	CTTGATGTCCTGACAGAGA	540
Qy	541	CACCCATCAACAGTATTAT	TTTCTCCATGAAGCGGTAC	GGGCGTGGGCGTGGAGCATC	600
Db	541	CACCCATCAACAGTATTAT	TTTCTCCATGAAGCGGTAC	GGGCGTGGGCGTGGAGCATC	600
Qy	601	TGTTGTCATTTGGGTACAC	GAGCAATCGCGCTGTAGT	AGCGGCCCATTAAGTTCCTCG	660
Db	601	TGTTGTCATTTGGGTACAC	GAGCAATCGCGCTGTAGT	AGCGGCCCATTAAGTTCCTCG	660
Qy	661	CGGCTCTGCTGCTGCTGCT	GGCATAAATATCTCATCG	CAATCAAAATTCAGCGGATAG	720
Db	661	CGGCTCTGCTGCTGCTGCT	GGCATAAATATCTCATCG	CAATCAAAATTCAGCGGATAG	720
Qy	721	CGGAACGGGAAGCGGACT	GAGTGCATGTCGGTTC	CAACAAACCATGCAAAATGCTGA	780
Db	721	CGGAACGGGAAGCGGACT	GAGTGCATGTCGGTTC	CAACAAACCATGCAAAATGCTGA	780

QY 781 ATGAGGCGATCGTTCCCACTGCGATGCTGGTTGCCAAGCATCAGATCGCGCTGGGCGCAA 840
DB 781 ATGAGGCGATCGTTCCCACTGCGATGCTGGTTGCCAAGCATCAGATCGCGCTGGGCGCAA 840
QY 841 TGGCGCCATTACCGAGTCCGGGCTCGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
DB 841 TGGCGCCATTACCGAGTCCGGGCTCGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAACCATCAACACGAGATTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAACCATCAACACGAGATTTC 960
QY 961 GCCTGCTGGGCGAAGACAGCTGAGCCGCTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
DB 961 GCCTGCTGGGCGAAGACAGCTGAGCCGCTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTTGGCCGCTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
DB 1021 AGGCAATCAGCTGTTGGCCGCTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
QY 1081 CGCAATCCGCTCTCCCGCGGTTGGCGGATTCAATTAATGCGCTGGCGACGAGTTT 1140
DB 1081 CGCAATCCGCTCTCCCGCGGTTGGCGGATTCAATTAATGCGCTGGCGACGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGTCAGTTAGCTCACTATTAG 1200
DB 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGTCAGTTAGCTCACTATTAG 1200
QY 1201 GCACAATTCATGTTTGACAGCTTATCATGACTGCAAGGTCGCAACCAATGCTTCGGCG 1260
DB 1201 GCACAATTCATGTTTGACAGCTTATCATGACTGCAAGGTCGCAACCAATGCTTCGGCG 1260
QY 1261 TCAGGCGCCATCGGAAGCTGGTATGCGTGTGCGAGTGTGTAATCACTGCGATTAATTCG 1320
DB 1261 TCAGGCGCCATCGGAAGCTGGTATGCGTGTGCGAGTGTGTAATCACTGCGATTAATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATATGTTTTTGGCGCGACATCAACGGTT 1380
DB 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATATGTTTTTGGCGCGACATCAACGGTT 1380
QY 1381 CTGGCAATATCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTTGTA 1440
DB 1381 CTGGCAATATCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTTGTA 1440
QY 1441 ATTGTAGCGGATAACAATTTACACAGGAAACAGCCAGTCCGTTAGGTGTTTTCAGGA 1500
DB 1441 ATTGTAGCGGATAACAATTTACACAGGAAACAGCCAGTCCGTTAGGTGTTTTCAGGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAGGTAAACTGGTAATCTGG 1560
DB 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAGGTAAACTGGTAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAAGTCGGTAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAATCGAAGAGAAATTCACAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAATCGAAGAGAAATTCACAGGTT 1680
QY 1681 GGGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACGACCGCTTGGTGGCTAC 1740
DB 1681 GGGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACGACCGCTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGGCTGTGGCTGAAATCACCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
DB 1741 GCTCAATCTGGGCTGTGGCTGAAATCACCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
DB 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAACCCCGCAAAACCTGGGAA 1920

DB 1861 GAAGCGTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAACCCCGCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGTAAAGAGCGCGCTGATGTTCAAC 1980
DB 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGTAAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTCGCTGAGCGGGGTTATGCGTTCAAG 2040
DB 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTCGCTGAGCGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAACGCGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
DB 2041 TATGAAAACGCGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACATGAAATGCAAGACCGGATTAC 2160
DB 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACATGAAATGCAAGACCGGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCGCTTTAATAAAGCGCAACAGCGATGACCATCAACCGGCCGTTG 2220
DB 2161 TCCATCGCAGAGCTGCGCTTTAATAAAGCGCAACAGCGATGACCATCAACCGGCCGTTG 2220
QY 2221 GCATGGTCCAACTCGACACAGCAAGTGAATTAATGTTGTAACGCTACTGCGACCTTC 2280
DB 2221 GCATGGTCCAACTCGACACAGCAAGTGAATTAATGTTGTAACGCTACTGCGACCTTC 2280
QY 2281 AAGGTCACCAATCCAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
DB 2281 AAGGTCACCAATCCAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTCTCTCGAAAACATCTGCTGACTGATGAAGGTCG 2400
DB 2341 CCGAACAAGAGCTGGCAAAAGAGTCTCTCGAAAACATCTGCTGACTGATGAAGGTCG 2400
QY 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGTAGCGCTGAACTCTTACGAGGAAG 2460
DB 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGTAGCGCTGAACTCTTACGAGGAAG 2460
QY 2461 TTGCGAAGAGTCCAAGTATTCGCGCCCAACATGGAACAAACGCCAGAAAGGTGAATCATG 2520
DB 2461 TTGCGAAGAGTCCAAGTATTCGCGCCCAACATGGAACAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACAATCCCGCAGATGTCGCGCTTTCTGGTATGCGCTGCTGAGTGTGATGATGATG 2580
DB 2521 CCGAACAATCCCGCAGATGTCGCGCTTTCTGGTATGCGCTGCTGAGTGTGATGATGATG 2580
QY 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
DB 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACCAACAACAATTAACAATAACAACCAACCTCGGGATCGAGGGAAGGATTTAGAAATTC 2700
DB 2641 AACCAACAACAATTAACAATAACAACCAACCTCGGGATCGAGGGAAGGATTTAGAAATTC 2700
QY 2701 GGATCCTCTTCTCTGTGGCCAGCGGCTCTGAGCCCGGGGAGAACCCCTATGCTGT 2760
DB 2701 GGATCCTCTTCTCTGTGGCCAGCGGCTCTGAGCCCGGGGAGAACCCCTATGCTGT 2760
QY 2761 CCGGAATGTTGTTAGTCTCTCAGCCAGAGAGCTCCCTGCTGCGCCACAGCGTACCCAC 2820
DB 2761 CCGGAATGTTGTTAGTCTCTCAGCCAGAGAGCTCCCTGCTGCGCCACAGCGTACCCAC 2820
QY 2821 ACGGTGAAAAACCGTATAAATGCGGCAAACTCTTTAGCCAGTCCAGCAAC 2880
DB 2821 ACGGTGAAAAACCGTATAAATGCGGCAAACTCTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGTTGCGCCATCAACGACTCATACTGCGGAGAGCCATCAAAATGTCCGAATGTGGC 2940
DB 2881 CTGTTGCGCCATCAACGACTCATACTGCGGAGAGCCATCAAAATGTCCGAATGTGGC 2940
QY 2941 AAGTCTTTTCCGAGTCCAGCAACTGTTGGCGCCACCAACGCTACTCACACCGGGAGAG 3000

```

Db 2941 AAGTCTTCAGTGATTGTCGTGATCTTCGAGGCGACCAACGTACTCACACCGGGGAGAG 3000
QY 3001 CCCTATGCTTGTCCGGAATGCTGAAGTCCTTCAGCACCGAGTGGCTCCTTGTGTTAGACAC 3060
Db 3001 CCCTATGCTTGTCCGGAATGCTGAAGTCCTTCTCAGAGCTCTCACCTGTGCGCCAC 3060
QY 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAAATGCCAGAGTGGCGGAAATCTTTTAGC 3120
QY 3121 CAGCGGGGCCACCTCGAAGCCCATCAAGCACTCATCTACTGGGAGAACCATACAAATGT 3180
Db 3121 CCGACGATAACCTGGTGGCCCATCAAGCACTCATCTACTGGGAGAACCATACAAATGT 3180
QY 3181 CCAGATGTGGCAAGTCTTCTCAACTTCAGGCAACTTGGTCCGTCACCAAGCTACTCAC 3240
Db 3181 CCAGATGTGGCAAGTCTTCTCAACTTCAGGCAACTTGGTCCGTCACCAAGCTACTCAC 3240
QY 3241 ACCGTTAAAAAACTAGTGGCGAGCGCGCGGCGGCGTACCGTACGAGTTCGGACTACGCT 3300
Db 3241 ACCGTTAAAAAACTAGTGGCGAGCGCGCGGCGGCGTACCGTACGAGTTCGGACTACGCT 3300

RESULT 3
AAD11588
ID AAD11588 standard; DNA; 3300 BP.
AC
XX
XX AAD11588;
DT 24-SEP-2001 (first entry)
DE Partial sequence of pMal-m1 and ZFPm1 DNA.
XX
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
KW modulation; plant technology; agriculture; ds.
OS Unidentified.
XX
FH Key
FT Location/Qualifiers
FT 2719..3270
FT /*tag= a
FT /product= "ZFPm1 protein"
FT /transl_except= (pos:2785..2787, aa:Asp)
FT /transl_except= (pos:2788..2790, aa:Pro)
FT /transl_except= (pos:2791..2793, aa:Gly)
FT /transl_except= (pos:2872..2874, aa:Arg)
FT /transl_except= (pos:2875..2877, aa:Ala)
FT /transl_except= (pos:2878..2880, aa:His)
FT /transl_except= (pos:2884..2886, aa:Glu)
FT /transl_except= (pos:2953..2955, aa:Gln)
FT /transl_except= (pos:2959..2961, aa:Ser)
FT /transl_except= (pos:3121..3123, aa:Arg)
FT /transl_except= (pos:3124..3126, aa:Ser)
FT /transl_except= (pos:3127..3129, aa:Asp)
FT /transl_except= (pos:3130..3132, aa:Asn)
FT /transl_except= (pos:3136..3138, aa:Val)
FT /transl_except= (pos:3139..3141, aa:Arg)
FT /transl_except= (pos:3205..3207, aa:Gln)
FT /transl_except= (pos:3208..3210, aa:Ala)
FT /transl_except= (pos:3211..3213, aa:Gly)
FT /transl_except= (pos:3214..3216, aa:His)
FT /transl_except= (pos:3220..3222, aa:Ala)
FT /transl_except= (pos:3223..3225, aa:Ser)
FT /note= "CDS does not include start and stop codon"
FT /partial
FT 2740..2790
FT /*tag= b
FT /bound_moiety= "F1-f2 primer"
FT 2770..2850
FT /*tag= c
FT /bound_moiety= "F1-f1 primer"
FT 2824..2889
FT /*tag= d

```

```

FT primer_bind
FT /bound_moiety= "F2-b primer"
FT 2867..2940
FT /*tag= e
FT /bound_moiety= "F2-f primer"
FT 2916..2973
FT /*tag= f
FT /bound_moiety= "F3-b1 primer"
FT 2953..3021
FT /*tag= g
FT /bound_moiety= "F3-b2 primer"
FT 2992..3042
FT /*tag= h
FT /bound_moiety= "F4-f2 primer"
FT 3022..3102
FT /*tag= i
FT /bound_moiety= "F4-f1 primer"
FT 3076..3141
FT /*tag= j
FT /bound_moiety= "F5-b primer"
FT 3119..3192
FT /*tag= k
FT /bound_moiety= "F5-f primer"
FT 3168..3225
FT /*tag= l
FT /bound_moiety= "F6-b1 primer"
FT 3205..3273
FT /*tag= m
FT /bound_moiety= "F6-b2 primer"
FT
XX WO200152620-A2.
PN
XX
XX 26-JUL-2001.
XX
XX 19-JAN-2001; 2001WO-US001817.
XX
PR 21-JAN-2000; 2000US-0177468P.
PR 21-JUL-2000; 2000US-00620897.
XX
XX (SCRI ) SCRIPPS RES INST.
XX (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
XX Barbas CF, Stege JT, Guan X, Dalmia B;
XX
XX WPI; 2001-465325/50.
XX P-PSDB; AAE06000.
XX
XX New zinc finger proteins, useful for modulating or regulating gene
XX expression and metabolic pathways in plants, e.g. for treating in the
XX plant cells a disorder that is associated with abnormal expression of the
XX target gene.
XX
XX Example 4; Page 138-139; 156pp; English.
XX
XX The patent discloses methods and compositions to modulate the expression
XX of a target gene in plant cells. The method involves providing plant
XX cells with a zinc finger protein (ZFP) which is capable of specifically
XX binding to a target nucleotide sequence or its complementary strand
XX within a target gene and allowing the ZFP binding to the target
XX nucleotide sequence, where the expression of the target gene in the plant
XX cells is modulated. The ZFP and fusions of the ZFP proteins are useful
XX for modulating or regulating gene expression and metabolic pathways in
XX plants. The ZFP, fusion proteins and methods are useful in plant and
XX agricultural technology. The method is useful particularly for treating a
XX disorder in the plant cells, where the disorder is associated with
XX abnormal expression of the target gene. The present DNA sequence is the
XX partial sequence of pMal-m1 and ZFPm1 DNA
XX
XX Sequence 3300 BP; 847 A; 874 C; 876 G; 703 T; 0 U; 0 Other;

```

```

Query Match          97.8%; Score 3226.4; DB 4; Length 3300;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3254; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

```

QY 1 CCGACACCATCGAATGCTGCAAAACCTTTTCGGGTATGCGCATGTAGCGCCGGAAGAGA 60
Db |||||
1 CCGACACCATCGAATGCTGCAAAACCTTTTCGGGTATGCGCATGTAGCGCCGGAAGAGA 60
QY 61 GTCAATTTCAGGGTGTGAATGTGAACACAGTAAACGTTATACGATGTCCGACAGTATGCCG 120
Db |||||
61 GTCAATTTCAGGGTGTGAATGTGAACACAGTAAACGTTATACGATGTCCGACAGTATGCCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCTCGCGTGTGTAACACGAGCCAGCCACGTTTCTGCGAAAA 180
Db |||||
121 GTGTCTCTTATCAGACCGTTTCCTCGCGTGTGTAACACGAGCCAGCCACGTTTCTGCGAAAA 180
QY 181 CGCGGGAAAAAGTGAACGCGGATGCGGAGCTGAATTTACATTTCCCAACCGCTGGGCAC 240
Db |||||
181 CGCGGGAAAAAGTGAACGCGGATGCGGAGCTGAATTTACATTTCCCAACCGCTGGGCAC 240
QY 241 AACAACTGCGCGGCAAAACAGTCGTTGCTGATTTGGCGTTTGCACCTTCCAGTCTGGCCCTGC 300
Db |||||
241 AACAACTGCGCGGCAAAACAGTCGTTGCTGATTTGGCGTTTGCACCTTCCAGTCTGGCCCTGC 300
QY 301 ACGGCGCTCGCAAAATTCGCGGCGATTAATCTCGCGCGATCAATCTGGGTCGACGC 360
Db |||||
301 ACGGCGCTCGCAAAATTCGCGGCGATTAATCTCGCGCGATCAATCTGGGTCGACGC 360
QY 361 TGGTGGTGTGATGTTAGAACGAGCGCGTGAAGCCGTGTAAGCGCGGTGCAACAATC 420
Db |||||
361 TGGTGGTGTGATGTTAGAACGAGCGCGTGAAGCCGTGTAAGCGCGGTGCAACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTCGCTGATGACACAGGATGCCA 480
Db |||||
421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTCGCTGATGACACAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCCTGCAATAATGTTCCGGCGTTATTTCTTGATGTTCTTGACCCAGA 540
Db |||||
481 TTGCTGTGGAAGCTGCCTGCAATAATGTTCCGGCGTTATTTCTTGATGTTCTTGACCCAGA 540
QY 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGGCGGTGAGACATC 600
Db |||||
541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGGCGGTGAGACATC 600
QY 601 TGGTCGCAATGGGTCAACAGCAAAATCGCGCTGTGTAGCGGGCCCAATTAAGTTCTGTCTCG 660
Db |||||
601 TGGTCGCAATGGGTCAACAGCAAAATCGCGCTGTGTAGCGGGCCCAATTAAGTTCTGTCTCG 660
QY 661 CGCGTCTCGCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db |||||
661 CGCGTCTCGCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGCGGACTGGAGTGCATGTCCGGTTTTCACAAACCATGCAAAATGCTGA 780
Db |||||
721 CGGAACGGGAAGCGGACTGGAGTGCATGTCCGGTTTTCACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTCCCACTCGATGCTGTGTCGAACAGATCAGATGCGCTGGCGCGAA 840
Db |||||
781 ATGAGGGCATCGTCCCACTCGATGCTGTGTCGAACAGATCAGATGCGCTGGCGCGAA 840
QY 841 TGC CGGCCATTAACCGAGTCCGGGCTGCGGTGTGGTGGGATATCTCGGTAGTGGGATAG 900
Db |||||
841 TGC CGGCCATTAACCGAGTCCGGGCTGCGGTGTGGTGGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCATCAACAGGATTTTC 960
Db |||||
901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCATCAACAGGATTTTC 960
QY 961 GCCTCTGGGGCAACACGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
Db |||||
961 GCCTCTGGGGCAACACGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAAACCAACCTGGCGGCCAATA 1080
Db |||||
1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAAACCAACCTGGCGGCCAATA 1080
QY 1081 CGCAAAACCGCCTCTCCCGCGCGCTTGGCCGATTCATTAATGCACTGGCAACAGAGTTT 1140

Db |||||
1081 CGCAAAACCGCCTCTCCCGCGCGTTGGCCGATTCATTAATGCACTGGCAGCAGAGTTT 1140
QY 1141 CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db |||||
1141 CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAATTCATGTTTGACAGCTTATCATCGACTGCAACGGTGCACCAATGCTTCTGGCG 1260
Db |||||
1201 GCACAATTCATGTTTGACAGCTTATCATCGACTGCAACGGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGCAGACCCATCGGAAGCTGTGATGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1320
Db |||||
1261 TCAGCAGACCCATCGGAAGCTGTGATGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1320
QY 1321 TGTCTGCTCAAGCGCAGCTCCCGTTCTGGATAATGTTTTTGGCGGACATCAATAACGGTT 1380
Db |||||
1321 TGTCTGCTCAAGCGCAGCTCCCGTTCTGGATAATGTTTTTGGCGGACATCAATAACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
Db |||||
1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
QY 1441 ATTGTGAGCGGATAACCAATTTTCACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
Db |||||
1441 ATTGTGAGCGGATAACCAATTTTCACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
QY 1501 GCACCTTACCACCAAGGACCAATAGATTATGAAAACTGAAGAGGTAAGTAACTGTTTTCGG 1560
Db |||||
1501 GCACCTTACCACCAAGGACCAATAGATTATGAAAACTGAAGAGGTAAGTAACTGTTTTCGG 1560
QY 1561 ATTTAACGCGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db |||||
1561 ATTTAACGCGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTACACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCGCCAGGTT 1680
Db |||||
1621 ACCGGAATTAAGTACACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCGCCAGGTT 1680
QY 1681 CGCGCAACTGCGCATGCGCCCTGACATATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db |||||
1681 CGCGCAACTGCGCATGCGCCCTGACATATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCGGAGCAAGCGTTCCAGGACAAGCTGTAT 1800
Db |||||
1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCGGAGCAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTCGGATGCGGTACGTTTACAAACGCAAGCTGATTTACCCGATCGCTGTT 1860
Db |||||
1801 CCGTTTACCTCGGATGCGGTACGTTTACAAACGCAAGCTGATTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
Db |||||
1861 GAAGCGTTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCTGATGTTCACAC 1980
Db |||||
1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCTGATGTTCACAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTTGGCCGCTGATTTGCTGCTGACGCGGGTTATGCGTTCAAG 2040
Db |||||
1981 CTGCAAGAACCGTACTTCACTTGGCCGCTGATTTGCTGCTGACGCGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAACAGCTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Db |||||
2041 TATGAAAAACGCAAGTACGACATTAACAGCTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGGTTGACCTGATTTAAAAACAAACATGAAATGCAAGACACCGATTAC 2160
Db |||||
2101 GGTCTGACCTTCTCGGTTGACCTGATTTAAAAACAAACATGAAATGCAAGACACCGATTAC 2160
QY 2161 TCCATCGCAAGAGCTGCGCTTTTAAAGCGGAAACAGCGATGACCATCAACCGCCCGTGG 2220
Db |||||

XX WPI, 2001-465325/50.
DR P-PSDB; AAE06002.
XX
PT New zinc finger proteins, useful for modulating or regulating gene
PT expression and metabolic pathways in plants, e.g. for treating in the
PT plant cells a disorder that is associated with abnormal expression of the
PT target gene.
XX
PS Example 4; Page 140-142; 156pp; English.
XX
CC The patent discloses methods and compositions to modulate the expression
CC of a target gene in plant cells. The method involves providing plant
CC cells with a zinc finger protein (ZFP) which is capable of specifically
CC binding to a target nucleotide sequence or its complementary strand
CC within a target gene and allowing the ZFP binding to the target
CC nucleotide sequence, where the expression of the target gene in the plant
CC cells is modulated. The ZFP and fusions of the ZFP proteins are useful
CC for modulating or regulating gene expression and metabolic pathways in
CC plants. The ZFP, fusion proteins and methods are useful in plant and
CC agricultural technology. The method is useful particularly for treating a
CC disorder in the plant cells, where the disorder is associated with
CC abnormal expression of the target gene. The present DNA sequence is the
CC partial sequence of pMal-m2 and ZFPm2 DNA
XX
SQ Sequence 3300 BP; 847 A; 874 C; 876 G; 703 T; 0 U; 0 Other;
Query Match 97.8%; Score 3226.4; DB 4; Length 3300;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3254; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 1 CGGACACATCGAATGGTGGCAAAACCTTTCGGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
DB 1 CGGACACCATCGAATGGTGGCAAAACCTTTCGGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTGAGGTGGTGAATGTGAACAGTAAACAGTATACGATGTCGAGAGTATGCG 120
DB 61 GTCAATTGAGGTGGTGAATGTGAACAGTAAACAGTATACGATGTCGAGAGTATGCG 120
QY 121 GTGTCTCTTATCAGACCGGTTTCCCGCGTGGTGAACAGGCGCACGCGTTTCTGGGAAAA 180
DB 121 GTGTCTCTTATCAGACCGGTTTCCCGCGTGGTGAACAGGCGCACGCGTTTCTGGGAAAA 180
QY 181 CGCGGAAAAAGTGAAGCGCGGATGGGAGCTGAAATTAATCCCAACCGCGTGGCAC 240
DB 181 CGCGGAAAAAGTGAAGCGCGGATGGGAGCTGAAATTAATCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAACAGTCTGTTGCTGATTTGGCGTTTCCACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGGCAACAGTCTGTTGCTGATTTGGCGTTTCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGGTGCGCAATTTGTGCGGCGGATTAATACTCTCGCGCGGATCAACTGGGTGCGAGCG 360
DB 301 ACGCGCGGTGCGCAATTTGTGCGGCGGATTAATACTCTCGCGCGGATCAACTGGGTGCGAGCG 360
QY 361 TGGTGGTGTGATGTAGAACGAAAGCGCGTGGAGCTGAAAGCGGCGGTGCAATTC 420
DB 361 TGGTGGTGTGATGTAGAACGAAAGCGCGTGGAGCTGAAAGCGGCGGTGCAATTC 420
QY 421 TTCTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACAGGATGCCA 480
DB 421 TTCTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCTCACTAATGTTCCGGCGGTTATTTCTGATGCTCTGACCCAGA 540
DB 481 TTGCTGTGGAAGCTGCTGCTCACTAATGTTCCGGCGGTTATTTCTGATGCTCTGACCCAGA 540
QY 541 CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGCGACTGGGCGGTGGAGCATC 600
DB 541 CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGCGACTGGGCGGTGGAGCATC 600
QY 601 TGGTGGCATTTGGGTACCGAATAATCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCTCGG 660
DB 601 TGGTGGCATTTGGGTACCGAATAATCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCTCGG 660

DB 601 TGGTGGCATTTGGGTACCGAATAATCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTGCGCTGCTGGCTGGGCATATAATATCTCATCTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGCGCTGCTGGCTGGGCATATAATATCTCATCTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAATTCGCTGA 780
DB 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAATTCGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTCGCATGCTGGTTGCGCAACGATCGCATGGCGTGGCGGCA 840
DB 781 ATGAGGGCATGTTCCCACTCGCATGCTGGTTGCGCAACGATCGCATGGCGTGGCGGCA 840
QY 841 TGGCGCCCATTAACGAGTCCGGGCTGCGCGTGGTGGCGATATCTCGGTAGTGGGATAG 900
DB 841 TGGCGCCCATTAACGAGTCCGGGCTGCGCGTGGTGGCGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCCATCAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGTCTGCACTCTCTCAGGCGCCAGGCGTGA 1020
DB 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGTCTGCACTCTCTCAGGCGCCAGGCGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCTCTCATCTGTTGAAAGAAAAACCAACCTGGCGCCCAATA 1080
DB 1021 AGGGCAATCAGCTGTTGCCCTCTCATCTGTTGAAAGAAAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGCAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGCTGGCAGCAGAGTTT 1140
DB 1081 CGCAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGCTGGCAGCAGAGTTT 1140
QY 1141 CCAGACTCGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTAGTCTCACTCATTAG 1200
DB 1141 CCAGACTCGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTAGTCTCACTCATTAG 1200
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACGCTGCAACCAATGTTCTGGCG 1260
DB 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACGCTGCAACCAATGTTCTGGCG 1260
QY 1261 TCAGGACCGCATCGGAAGCTGTTGATGCTGTCGAGCTGTAATCATCTGATTAATTCG 1320
DB 1261 TCAGGACCGCATCGGAAGCTGTTGATGCTGTCGAGCTGTAATCATCTGATTAATTCG 1320
QY 1321 TGTGCTCAAGCGCACCTCCCGTCTTGATTAATGTTTTTGGCGCGACATCAATAACGGTT 1380
DB 1321 TGTGCTCAAGCGCACCTCCCGTCTTGATTAATGTTTTTGGCGCGACATCAATAACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
DB 1381 CTGGCAAAATTTCTGAAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
QY 1441 ATTGTAGCGGATTAACAAATTTCAACAGGAAACAGCCAGTCCGTTTAGTGTGTTTCAAGA 1500
DB 1441 ATTGTAGCGGATTAACAAATTTCAACAGGAAACAGCCAGTCCGTTTAGTGTGTTTCAAGA 1500
QY 1501 GCACCTTCAACCAACAGGACCATAGATTAATGAATGAATGAAGAGGTAATACTGGTAACTCG 1560
DB 1501 GCACCTTCAACCAACAGGACCATAGATTAATGAATGAATGAAGAGGTAATACTGGTAACTCG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACCGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTAACGGCGATAAAGGCTATAACCGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCCTCCAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCCTCCAGGTT 1680
QY 1681 GGGCAATCTGGCGATGGCCCTGACATTAATCTCTGGGCAACGACCGCTTTGTTGGCTAC 1740
DB 1681 GGGCAATCTGGCGATGGCCCTGACATTAATCTCTGGGCAACGACCGCTTTGTTGGCTAC 1740

```
QY 1741 GCTCAATCTGGCTGTGTGGCTGAATTCACCCCGGACAAAGCGTTCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGTGGCTGAATTCACCCCGGACAAAGCGTTCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACTCGGATGCCGTAGCTTTACAAAGCGCAAGCTGATTGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACTCGGATGCCGTAGCTTTACAAAGCGCAAGCTGATTGCTTACCCGATCGCTGT 1860
QY 1861 GAAAGCGTTATCGCTGATTTATAAACAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAAGCGTTATCGCTGATTTATAAACAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTACCTGGCGCGCTGATTGCTGCTGACGGGGTATTGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTGGCGCGCTGATTGCTGCTGACGGGGTATTGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGGTGGATTAACGCTGGCGGCAAGACG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGGTGGATTAACGCTGGCGGCAAGACG 2100
QY 2101 GGTCTGACCTTCTGGTTGACTGATTAAGAACTGAATTAAGAACTGAATTAAGAACTGAAT 2160
Db 2101 GGTCTGACCTTCTGGTTGACTGATTAAGAACTGAATTAAGAACTGAATTAAGAACTGAAT 2160
QY 2161 TCCATCGCAGAGCTGCTTTAAATAAGCGCAACACAGCGATGACCATCAACGGCCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAAATAAGCGCAACACAGCGATGACCATCAACGGCCCGCTGG 2220
QY 2221 GCATGGTCCAAATCGACACACAGCAAAAGTGAATTAAGCGTAAAGCGTACTCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACACAGCAAAAGTGAATTAAGCGTAAAGCGTACTCCGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAAGCGCGCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAAGCGCGCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCCGAAACTATCTGCTGACTGATGAAGGTCGTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCCGAAACTATCTGCTGACTGATGAAGGTCGTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGTCCGCTAGCGCTGAAAGTCTTACGAGGAAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTCCGCTAGCGCTGAAAGTCTTACGAGGAAAG 2460
QY 2461 TTGGCGAAAGATCCACGTATTGCGCCACCATGGAACACGCCGAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTATTGCGCCACCATGGAACACGCCGAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCGCGAGATGTCGCTTTCTGTTATGCCGTGCGTACTGCGGTGATCAAGCC 2580
Db 2521 CCGAACATCCGCGAGATGTCGCTTTCTGTTATGCCGTGCGTACTGCGGTGATCAAGCC 2580
QY 2581 GCCAGCGTGTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTGTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGGGGAGAGGATTCAGAAATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGGGGAGAGGATTCAGAAATTC 2700
QY 2701 GGATCTCTTCTCTGTGGCCAGCGCGCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTGGCCAGCGCGCTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGCTGTAAGTCTTTCAGCCAGACAGCTCCCTGCTGTCGCCACAGCGTACCCAC 2820
Db 2761 CCGGAATGCTGTAAGTCTTTCAGCCAGACAGCTCCCTGCTGTCGCCACAGCGTACCCAC 2820
```

RESULT 5

```
AAD11590
ID AAD11590 standard; DNA; 3300 BP.
XX
AC AAD11590;
XX
DT 24-SEP-2001 (first entry)
XX
DE Partial sequence of pMal-m3 and ZFPm3 DNA.
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
KW modulation; plant technology; agriculture; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 2719..3270
FT /tag= a
FT /product= "ZFPm3 protein"
FT /note= "CDS does not include start and stop codon"
FT primer_bind 2740..2790
FT /tag= b
FT primer_bind 2770..2850
FT /tag= c
FT primer_bind 2824..2889
FT /tag= d
FT primer_bind 2867..2940
FT /tag= e
FT primer_bind 2916..2973
FT /tag= f
FT primer_bind 2953..3021
FT /tag= g
FT primer_bind 2992..3042
```

FT primer_bind /*tag= h
FT /bound_moiety= "F4-f2 primer"
FT 3022..3102
FT /*tag= i
FT /bound_moiety= "F4-f1 primer"
FT 3076..3141
FT /*tag= j
FT /bound_moiety= "F5-b primer"
FT 3119..3192
FT /*tag= k
FT /bound_moiety= "F5-f primer"
FT 3168..3225
FT /*tag= l
FT /bound_moiety= "F6-b1 primer"
FT 3205..3273
FT /*tag= m
FT /bound_moiety= "F6-b2 primer"
XX WO200152620-A2.
XX
XX
XX 26-JUL-2001.
XX
XX 19-JAN-2001; 2001WO-US001817.
XX
XX 21-JAN-2000; 2000US-0177468P.
XX 21-JUL-2000; 2000US-00620897.
XX
XX (SCRI) SCRIPPS RES INST.
XX (SYCN) SYNGENTA AGRIC DISCOVERY INC.
XX
XX Barbas CF, Stege JT, Guan X, Dalmia B;
XX WPI; 2001-465325/50.
XX P-PSDB; AAE06003.
XX
XX New zinc finger proteins, useful for modulating or regulating gene
XX expression and metabolic pathways in plants, e.g. for treating in the
XX plant cells a disorder that is associated with abnormal expression of the
XX target gene.
XX
XX Example 4; Page 143-145; 156pp; English.
XX
XX The patent discloses methods and compositions to modulate the expression
XX of a target gene in plant cells. The method involves providing plant
XX cells with a zinc finger protein (ZFP) which is capable of specifically
XX binding to a target nucleotide sequence or its complementary strand
XX within a target gene and allowing the ZFP binding to the target
XX nucleotide sequence, where the expression of the target gene in the plant
XX cells is modulated. The ZFP and fusions of the ZFP proteins are useful
XX for modulating or regulating gene expression and metabolic pathways in
XX plants. The ZFP, fusion proteins and methods are useful in plant and
XX agricultural technology. The method is useful particularly for treating a
XX disorder in the plant cells, where the disorder is associated with
XX abnormal expression of the target gene. The present DNA sequence is the
XX Partial sequence of pMal-m3 and ZFPm3 DNA
SQ Sequence 3300 BP; 845 A; 878 C; 877 G; 700 T; 0 U; 0 Other;

Query Match 97.5%; Score 3216.8; DB 4; Length 3300;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 3248; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
DB 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60

QY 61 GTCAATTACGGTGGTGAATGTGAACCAACGTAACGTTATACGATTCGCAGAGTATGCCG 120
DB 61 GTCAATTACGGTGGTGAATGTGAACCAACGTAACGTTATACGATTCGCAGAGTATGCCG 120

QY 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACCAACGTCAGCCACCGTTTCTCGGAAA 180
DB 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACCAACGTCAGCCACCGTTTCTCGGAAA 180

QY 181 CGCGGAAAAAGTGGAAAGCGCGCATGGCGGAGCTGAATTACATTCCTCAACCGCTGGCAC 240
DB 181 CGCGGAAAAAGTGGAAAGCGCGCATGGCGGAGCTGAATTACATTCCTCAACCGCTGGCAC 240

QY 241 AACAACTGGCGGGCAAAACAGTCTGTGTTGCTGATTGGCGTTGCCACCTTCAGTCTGGGCCCTGCG 300
DB 241 AACAACTGGCGGGCAAAACAGTCTGTGTTGCTGANTGGCGTTGCCACCTTCAGTCTGGGCCCTGCG 300

QY 301 ACGCCCGCTCGCAAAATTGTTCGGCGCGATTAAATCTCGCGCGGATCAATCACTGGGTGCGCAGCG 360
DB 301 ACGCCCGCTCGCAAAATTGTTCGGCGCGATTAAATCTCGCGCGGATCAATCACTGGGTGCGCAGCG 360

QY 361 TGGTGGTGTTCATGGTAGAACGAAGCGCGCTGCAAGCTGTAAAGCGCGGTGCGACAATC 420
DB 361 TGGTGGTGTTCATGGTAGAACGAAGCGCGCTGCAAGCTGTAAAGCGCGGTGCGACAATC 420

QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGATGATCAACCAAGGATGCCA 480
DB 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGATGATCAACCAAGGATGCCA 480

QY 481 TTGCTGTGGAAGCTCCCTGCACTAATGTTCCGGCGTTAATTTCTTGATGTTCTCTGACCCAGA 540
DB 481 TTGCTGTGGAAGCTCCCTGCACTAATGTTCCGGCGTTAATTTCTTGATGTTCTCTGACCCAGA 540

QY 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGGCACTGGCGCTGGAGCATC 600
DB 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGGCACTGGCGCTGGAGCATC 600

QY 601 TGGTCGCATTGCGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAAGTTCTGTCTCGG 660
DB 601 TGGTCGCATTGCGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAAGTTCTGTCTCGG 660

QY 661 CGCGTCTCGCTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTCGCTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720

QY 721 CGGAACGGGAAGGACCTGGAGTGCATGTCGGTTTTCACAAACCAATGCAAAATGCTGA 780
DB 721 CGGAACGGGAAGGACCTGGAGTGCATGTCGGTTTTCACAAACCAATGCAAAATGCTGA 780

QY 781 ATGAGGGCATGTTCCCACTCGCATGCGTGTGTTGCAACGATCAGATGCGCTGGCGGCA 840
DB 781 ATGAGGGCATGTTCCCACTCGCATGCGTGTGTTGCAACGATCAGATGCGCTGGCGGCA 840

QY 841 TGGCGCCATTACCGAGTCCCGGCTGCGGTTGGTGGCGGATATCTCGGTAGTGGGATAG 900
DB 841 TGGCGCCATTACCGAGTCCCGGCTGCGGTTGGTGGCGGATATCTCGGTAGTGGGATAG 900

QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAATCAAAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAATCAAAACAGGATTTTC 960

QY 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
DB 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020

QY 1021 AGGCAATCAGCTGTTGCGCGCTCTCACTGGTGAAGAAACCAACCACTGGCGCCCAATA 1080
DB 1021 AGGCAATCAGCTGTTGCGCGCTCTCACTGGTGAAGAAACCAACCACTGGCGCCCAATA 1080

QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCCGATTCAATTAATGAGTGGACGACAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGGTTGGCCGATTCAATTAATGAGTGGACGACAGGTTT 1140

QY 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACCAATTAATGAGTGGACGACAGGTTT 1200
DB 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACCAATTAATGAGTGGACGACAGGTTT 1200

QY 1201 GCACAATTTCTCATGTTTTCAGACGCTTATCATGCACTGCAAGCGGTGCAACCAATGTTCTGGCG 1260
DB 1201 GCACAATTTCTCATGTTTTCAGACGCTTATCATGCACTGCAAGCGGTGCAACCAATGTTCTGGCG 1260

QY	1261	TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGAGAGTCTGTAATCACTGATTAATTCG	1320
Db	1261	TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGAGAGTCTGTAATCACTGATTAATTCG	1320
QY	1321	TGTCGCTCAAGCGGACATCCCGTCTGAGTATGTTTGGCCGAGCATATAACGGTT	1380
Db	1321	TGTCGCTCAAGCGGACATCCCGTCTGAGTATGTTTGGCCGAGCATATAACGGTT	1380
QY	1381	CTGGCAATATTTCTGAATAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA	1440
Db	1381	CTGGCAATATTTCTGAATAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA	1440
QY	1441	ATTGTAGCGGATTAACAATTTACACAGAAACAGCCAGTCCGTTAGGTTTTCACGA	1500
Db	1441	ATTGTAGCGGATTAACAATTTACACAGAAACAGCCAGTCCGTTAGGTTTTCACGA	1500
QY	1501	GCACCTTCAACCAAGGACCATAGATTATGAACCTGAAGAGTAAACTGGTAACTCTGG	1560
Db	1501	GCACCTTCAACCAAGGACCATAGATTATGAACCTGAAGAGTAAACTGGTAACTCTGG	1560
QY	1561	ATTAAACGCGATTAAGGCTATAACCGTCTCGCTGAAGTCCGTAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAAACGCGATTAAGGCTATAACCGTCTCGCTGAAGTCCGTAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATTAACCTGGAGAGAAATTCGCCAGGTT	1680
Db	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATTAACCTGGAGAGAAATTCGCCAGGTT	1680
QY	1681	CGCGCAACTGGCGATGGCCCTGACATTAATCTTGGGCACACGCGCTTTGGTGGCTAC	1740
Db	1681	CGCGCAACTGGCGATGGCCCTGACATTAATCTTGGGCACACGCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCCTGTGGCTGAATCAACCCCGAACAGGTTCCAGGACAAAGCTGTAT	1800
Db	1741	GCTCAATCTGGCCTGTGGCTGAATCAACCCCGAACAGGTTCCAGGACAAAGCTGTAT	1800
QY	1801	CCGTTTACCTGGATGGCTGACGTTACAAACGCGAAGCTGATTTACCGATCGCTGTT	1860
Db	1801	CCGTTTACCTGGATGGCTGACGTTACAAACGCGAAGCTGATTTACCGATCGCTGTT	1860
QY	1861	GAAGCTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCCGCAAAACCTGGGAA	1920
Db	1861	GAAGCTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCCGCAAAACCTGGGAA	1920
QY	1921	GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGTAAAGCGCGCTGATTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGTAAAGCGCGCTGATTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGTCGAGCGGGGTTATCGTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGTCGAGCGGGGTTATCGTTCAAG	2040
QY	2041	TATGAAACCGCAAGTACGACATTTAAAGACGTGGCGCTGGATACGCTGGCGGAAAGCG	2100
Db	2041	TATGAAACCGCAAGTACGACATTTAAAGACGTGGCGCTGGATACGCTGGCGGAAAGCG	2100
QY	2101	GCTGACCTTCTGCTGATTAAGAACTGAAGCGAAAGTAAAGCGCGCTGATTTCAAC	2160
Db	2101	GCTGACCTTCTGCTGATTAAGAACTGAAGCGAAAGTAAAGCGCGCTGATTTCAAC	2160
QY	2161	TCCATCCGAGAGCTGCTTTTAAAGAGCGAAGACGCGATGACCATCAACGCGCGTGG	2220
Db	2161	TCCATCCGAGAGCTGCTTTTAAAGAGCGAAGACGCGATGACCATCAACGCGCGTGG	2220
QY	2221	GCATGCTCAACATCGACACAGCAAGTGAATTTATGTTAAGCGTACTTCCGACCTTC	2280
Db	2221	GCATGCTCAACATCGACACAGCAAGTGAATTTATGTTAAGCGTACTTCCGACCTTC	2280
QY	2281	AAGGTCACATCCCAACCGTTGTTGGCTGCTGAGCGCAGGTATTAACGCGCGCAGT	2340
Db	2281	AAGGTCACATCCCAACCGTTGTTGGCTGCTGAGCGCAGGTATTAACGCGCGCAGT	2340
QY	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTCCGAAAACTATCTGCTGACTGATGAAGGCTG	2400

Db	2341	CCGAACAAAGAGCTGGCAAAAGAGTTCTCCGAAAACTATCTGCTGACTGATGAAGGCTG	2400
QY	2401	GAAGCGGTTAATAAGACAAACCGCTGGTCCCGTAGCTGAAAGTCTTACGAGAAAG	2460
Db	2401	GAAGCGGTTAATAAGACAAACCGCTGGTCCCGTAGCTGAAAGTCTTACGAGAAAG	2460
QY	2461	TTGCGGAAAGATCCACGATTTGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG	2520
Db	2461	TTGCGGAAAGATCCACGATTTGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG	2520
QY	2521	CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCTGCTGCTGCTGCTGCTG	2580
Db	2521	CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCTGCTGCTGCTGCTGCTG	2580
QY	2581	GCAGCGGCTGCTGAGTCTGATGAAGCCTGAAAGACGCGCAGACTAATTCGAGCTG	2640
Db	2581	GCAGCGGCTGCTGAGTCTGATGAAGCCTGAAAGACGCGCAGACTAATTCGAGCTG	2640
QY	2641	AACAAACAAACAAATTAACAACCTCGGATCGAGGAGAGGATTCAGAAATTC	2700
Db	2641	AACAAACAAACAAATTAACAACCTCGGATCGAGGAGAGGATTCAGAAATTC	2700
QY	2701	GGATCTCTTCTGCTGCTGCGCCAGCGGCTCGAGCCCGGGGAGAGCCCTATGCTGT	2760
Db	2701	GGATCTCTTCTGCTGCTGCGCCAGCGGCTCGAGCCCGGGGAGAGCCCTATGCTGT	2760
QY	2761	CCGGAATGCTGTAAGTCTTTCAGCGAGCAGCTCCTGCTGCTGCTGCTGCTGCTGCTG	2820
Db	2761	CCGGAATGCTGTAAGTCTTTCAGCGAGCAGCTCCTGCTGCTGCTGCTGCTGCTGCTG	2820
QY	2821	ACGGGTGAAAAACCGTATAAATCCAGAGTGGCGGAAATCTTTTAGCCAGTCCAGCA	2880
Db	2821	ACGGGTGAAAAACCGTATAAATCCAGAGTGGCGGAAATCTTTTAGCCAGTCCAGCA	2880
QY	2881	CTGGTGGCCATCAACGCACTACTTGGCGAGAGCCATACAAATGTCAGAAATGTTGC	2940
Db	2881	CTGGTGGCCATCAACGCACTACTTGGCGAGAGCCATACAAATGTCAGAAATGTTGC	2940
QY	2941	AAGTCTTTCAGCGAGTCCAGCACTGCTGCGCCACCAAGCTACTCACACCGGGGAGA	3000
Db	2941	AAGTCTTTCAGCGAGTCCAGCACTGCTGCGCCACCAAGCTACTCACACCGGGGAGA	3000
QY	3001	CCCTATGCTTGTCCGGAATGTTGTAAGTCTTTCAGCAACAGTGGCTCTTGGTTAGAC	3060
Db	3001	CCCTATGCTTGTCCGGAATGTTGTAAGTCTTTCAGCAACAGTGGCTCTTGGTTAGAC	3060
QY	3061	CAGCGTACCCACACCGGGTGAACCGTATAATGTCAGAGTGGCGGAAATCTTTTAGT	3120
Db	3061	CAGCGTACCCACACCGGGTGAACCGTATAATGTCAGAGTGGCGGAAATCTTTTAGT	3120
QY	3121	CAGCGCGCCACCTGGAACGCGCATCAACGCACTTCACTTGGCGGAGAGCCATCAATGT	3180
Db	3121	CAGCGCGCCACCTGGAACGCGCATCAACGCACTTCACTTGGCGGAGAGCCATCAATGT	3180
QY	3181	CCAGAAATGTCGCAAGTCTTTCCTCAACTTCAGGCACTTGGTCCGTCACCAAGTACTC	3240
Db	3181	CCAGAAATGTCGCAAGTCTTTCCTCAACTTCAGGCACTTGGTCCGTCACCAAGTACTC	3240
QY	3241	ACCGGTAAAAAACTAGTGGCGCAGCGCGGCTAGTCCCGTACGAGTTCGCGACTACGT	3300
Db	3241	ACCGGTAAAAAACTAGTGGCGCAGCGCGGCTAGTCCCGTACGAGTTCGCGACTACGT	3300

RESULT 6

AAH25681

ID AAH25681 standard; DNA; 8101 BP.

XX

AC AAH25681;

XX

DT 05-SEP-2001 (first entry)

XX

DE Nucleotide sequence of human Cdc25C phosphatase DNA..

XX Fusion protein; maltose binding protein; MBP; Cdc25C phosphatase;
KW Cdc25B1; Cdc25B2; Cdc25B3; cyclin-dependent kinase; CDK; cell division;
KW ss.
XX
XX Homo sapiens.
XX WO200144467-A2.
XX
XX PD 21-JUN-2001.
XX
XX PF 13-DEC-2000; 2000WO-FR003496.
XX
XX PR 14-DEC-1999; 99FR-00015722.
XX PR 30-MAY-2000; 2000FR-00006883.
XX PR 21-SEP-2000; 2000FR-00012008.
XX
XX PA (SCRC) SCRAS SOC CONSEILS RECH & APPL SCI.
XX
XX PI Goubin-Gramatica F, Ducommun B, Prevost G;
XX
XX DR WPI; 2001-398152/42.
XX
XX PT New fusion protein of maltose-binding protein and Cdc phosphatase, useful
XX for identifying phosphatase modulators for regulating the cell cycle.
XX
XX PS Claim 7; Page 9-13; 56pp; French.
XX
XX The present sequence encodes human Cdc25C phosphatase, and was expressed
XX in Escherichia coli. It is used to make fusion proteins with Escherichia
XX coli maltose binding protein (MBP). The specification describes fusion
XX proteins of MBP and one of Cdc 25B1, 25B2, 25B3 and 25C. Cdc phosphatases
XX are involved in activation of cyclin-dependent kinases (CDK) that control
XX cell division. The fusion proteins are used to identify modulators of the
XX specified human Cdc phosphatases, potentially useful for regulation of
XX cell division. They are also for studying physiological or
XX physiopathological activities of Cdc phosphatase
XX
XX SQ Sequence 8101 BP; 2127 A; 1987 C; 2121 G; 1866 T; 0 U; 0 Other;
Query Match 82.0%; Score 2705.8; DB 4; Length 8101;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CGACACCATCGAATGGTGCAAAACCTTTCCGGTATGCGCATGATAGCGCCCGGAAGA 60
DB 1 CGACACCATCGAATGGTGCAAAACCTTTCCGGTATGCGCATGATAGCGCCCGGAAGA 60
QY 61 GTCATTCAGGGTGTGAATGTGAACCAACGTAACGTTATACGATGTCCGAGATGCGG 120
DB 61 GTCATTCAGGGTGTGAATGTGAACCAACGTAACGTTATACGATGTCCGAGATGCGG 120
QY 121 GTGTCTTATCAGACCGTTTCCCGGTGTGCAACAGCCAGCCAGCTTTCTGCGAATA 180
DB 121 GTGTCTTATCAGACCGTTTCCCGGTGTGGAACAGCCAGCCAGCTTTCTGCGAATA 180
QY 181 CGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTTACATTCCTCAACCGCGTGGC 240
DB 181 CGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTTACATTCCTCAACCGCGTGGC 240
QY 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTGGCCACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTGGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCCCGTGGCAAAATGTTCGGCGGATTTAAATCTCGGCCGATCAACTGGGTGCCAGG 360
DB 301 ACGCCCGTGGCAAAATGTTCGGCGGATTTAAATCTCGGCCGATCAACTGGGTGCCAGG 360
QY 361 TGGTGTGTGCGATGGTAGAACGAAGCGCGTCCAAAGCCTGTAAAGCGGCGGTGCAAAATC 420
DB 361 TGGTGTGTGCGATGGTAGAACGAAGCGCGTCCAAAGCCTGTAAAGCGGCGGTGCAAAATC 420
QY 421 TTCTCGGCAACGCGTCTAGTGGGCTGATCAITAACTATCCGCTGGATGACAGGATGCCA 480

DB 421 TTCTCGGCAACGCGTCTAGTGGGCTGATCAITAACTATCCGCTGGATGACAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTCCCTGCACTAATGTTCCGGCGTTATTTCTTGATCTCTCTGACCA 540
DB 481 TTGCTGTGGAAGCTCCCTGCACTAATGTTCCGGCGTTATTTCTTGATCTCTCTGACCA 540
QY 541 CACCCATCAACAGTATTTCTCCCATGAAGAGGTACGGACTGGCGGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTTCTCCCATGAAGAGGTACGGACTGGCGGTGGAGCATC 600
QY 601 TGGTCGCAATTCGGTCACCAAGCAATTCGGCTGTAGCGGGCCCATTAAGTTCTGCTCGG 660
DB 601 TGGTCGCAATTCGGTCACCAAGCAATTCGGCTGTAGCGGGCCCATTAAGTTCTGCTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTTTCAACAAACCATGCAATGCTGA 780
DB 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTTTTCAACAAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGCCAAGATCAGATGCGCTGGCGCA 840
DB 781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGCCAAGATCAGATGCGCTGGCGCA 840
QY 841 TGGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGTAGTGGGATAG 900
DB 841 TGGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGTAGTGGGATAG 900
QY 901 AGATACCGAAGACAGTCATGTTATATCCCGCGTTAAACCAATCAACAGGATTTTC 960
DB 901 AGATACCGAAGACAGTCATGTTATATCCCGCGTTAAACCAATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGTCGAACCTCTCAGGCGCCAGGCGTGA 1020
DB 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGTCGAACCTCTCAGGCGCCAGGCGTGA 1020
QY 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAAAACCACTTGGCGCCCAATA 1080
DB 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAAGAAAAACCACTTGGCGCCCAATA 1080
QY 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGTGGCAACAGGTTT 1140
DB 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGTGGCAACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTCACTATTAG 1200
DB 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTCACTATTAG 1200
QY 1201 GCACAAATCTCATGTTTGACAGCTTTATCATCGACTGCAAGGTGCAACCAATGCTTGGCG 1260
DB 1201 GCACAAATCTCATGTTTGACAGCTTTATCATCGACTGCAAGGTGCAACCAATGCTTGGCG 1260
QY 1261 TCAGCAGCCATCGGAAGCTGTGTTATGCGTGTGAGGTGTAATCACTGCAATTAATTCG 1320
DB 1261 TCAGCAGCCATCGGAAGCTGTGTTATGCGTGTGAGGTGTAATCACTGCAATTAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACTCCCGTTCTGATTAATGTTTTTGGCGGACATCAATACGTT 1380
DB 1321 TGTGCTCAAGCGGCACTCCCGTTCTGATTAATGTTTTTGGCGGACATCAATACGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGCACAATTAATCATCGGCTCGTATATGTTGA 1440
DB 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGCACAATTAATCATCGGCTCGTATATGTTGA 1440
QY 1441 ATTGTAGCGGATAACAATTTTCAACAGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500
DB 1441 ATTGTAGCGGATAACAATTTTCAACAGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAACCTGAAGAGGTAAATCTGTAATCTCG 1560

Db	1501	GCAC	TTCA	CCAA	CA	AGG	CCATAG	ATTATG	AAAA	TCGA	AAAGGTA	AACTCG	TATCTGG	1560		
Qy	1561	ATTAA	CGGCGA	TAAAG	GGCTAT	AAC	GGTCTCG	CTGA	AGTCGG	TAGAA	ATTCG	AGAAAGAT	1620			
Db	1561	ATTAA	CGGCGA	TAAAG	GGCTAT	AAC	GGTCTCG	TGA	AGTCGG	TAGAA	ATTCG	AGAAAGAT	1620			
Qy	1621	ACCG	GAATTA	AAGT	CAC	GGTTAG	GCAT	CCGG	ATAACT	TGG	AAGAGAA	ATTC	CCCA	CAGGTT	1680	
Db	1621	ACCG	GAATTA	AAGT	CAC	GGTTAG	GCAT	CCGG	ATAACT	TGG	AAGAGAA	ATTC	CCCA	CAGGTT	1680	
Qy	1681	GCGG	CAACTGG	CGAT	TGG	CCCTG	ACAT	TATCT	TGGG	CAC	AGCCG	CTT	TGGT	GGCTAC	1740	
Db	1681	GCGG	CAACTGG	CGAT	TGG	CCCTG	ACAT	TATCT	TGGG	CAC	AGCCG	CTT	TGGT	GGCTAC	1740	
Qy	1741	GCTCA	ATCTGG	CCCTG	TTGG	CTGA	ATACT	CCCC	CGGAC	AAAG	CGTTCC	AGGCA	CAAG	CTGTAT	1800	
Db	1741	GCTCA	ATCTGG	CCCTG	TTGG	CTGA	ATACT	CCCC	CGGAC	AAAG	CGTTCC	AGGCA	CAAG	CTGTAT	1800	
Qy	1801	CGG	TTTACT	CTGG	ATG	CGT	AGTTT	AC	ACG	CAAG	CTG	ATG	CTT	ACCG	ATCG	1860
Db	1801	CGG	TTTACT	CTGG	ATG	CGT	AGTTT	AC	ACG	CAAG	CTG	ATG	CTT	ACCG	ATCG	1860
Qy	1861	GAAG	CGTTAT	CG	CTGA	TTTAA	CA	AAAG	ATCTG	TCG	CGAA	ACCC	CGC	CA	AAAA	1920
Db	1861	GAAG	CGTTAT	CG	CTGA	TTTAA	CA	AAAG	ATCTG	TCG	CGAA	ACCC	CGC	CA	AAAA	1920
Qy	1921	GAGAT	CCCG	CGCTGG	ATA	AGAACT	GA	AGCG	AAAG	AGT	AG	AGCG	CG	CTG	ATG	1980
Db	1921	GAGAT	CCCG	CGCTGG	ATA	AGAACT	GA	AGCG	AAAG	AGT	AG	AGCG	CG	CTG	ATG	1980
Qy	1981	CTG	CAAGAAC	CGT	ACT	CT	CAC	CTGG	CCG	CTG	AT	CTCT	GTG	ACG	GGGG	1040
Db	1981	CTG	CAAGAAC	CGT	ACT	CT	CAC	CTGG	CCG	CTG	AT	CTCT	GTG	ACG	GGGG	1040
Qy	2041	TATG	AAACG	CGA	AGT	ACG	ATTA	A	AG	ACG	TGG	GG	CTG	GAT	AAC	2100
Db	2041	TATG	AAACG	CGA	AGT	ACG	ATTA	A	AG	ACG	TGG	GG	CTG	GAT	AAC	2100
Qy	2101	GGT	CTG	AC	CTTCT	CTG	TTTG	AC	CTG	ATTA	AAAA	CA	CA	CA	CA	2160
Db	2101	GGT	CTG	AC	CTTCT	CTG	TTTG	AC	CTG	ATTA	AAAA	CA	CA	CA	CA	2160
Qy	2161	TCCAT	CG	CAG	AAG	CTG	CTTTA	T	AAAG	GCG	AAAC	AG	CG	AT	AAC	2220
Db	2161	TCCAT	CG	CAG	AAG	CTG	CTTTA	T	AAAG	GCG	AAAC	AG	CG	AT	AAC	2220
Qy	2221	GCAT	GGT	CCAA	CTCG	AC	CAAG	CA	AGT	AGT	AT	TGGT	TA	AGG	CT	2280
Db	2221	GCAT	GGT	CCAA	CTCG	AC	CAAG	CA	AGT	AGT	AT	TGGT	TA	AGG	CT	2280
Qy	2281	AAG	GGTCA	AC	ATCC	AA	ACCG	TT	CG	TTGG	CG	CTG	AG	CG	CAG	2340
Db	2281	AAG	GGTCA	AC	ATCC	AA	ACCG	TT	CG	TTGG	CG	CTG	AG	CG	CAG	2340
Qy	2341	CCG	AA	CA	AG	AG	CTG	CG	AAAG	AGT	CT	CTCG	AA	AA	CT	2400
Db	2341	CCG	AA	CA	AG	AG	CTG	CG	AAAG	AGT	CT	CTCG	AA	AA	CT	2400
Qy	2401	GAAG	CGGTTA	TAA	AG	CA	AA	CCG	TTTCTG	TG	CG	CTAG	CG	CTGA	AG	2460
Db	2401	GAAG	CGGTTA	TAA	AG	CA	AA	CCG	TTTCTG	TG	CG	CTAG	CG	CTGA	AG	2460
Qy	2461	TTGG	CG	AAAG	ATCC	AC	GTAT	TG	CG	CC	CA	CCAT	GG	AAAA	CG	2520
Db	2461	TTGG	CG	AAAG	ATCC	AC	GTAT	TG	CG	CC	CA	CCAT	GG	AAAA	CG	2520
Qy	2521	CCG	AA	CA	CTCC	G	AG	ATG	CCG	TTTCTG	TG	AT	TG	CG		

Query Match 82.0%; Score 2705.8; DB 6; Length 8101; Best Local Similarity 99.9%; Pred. No. 0; Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;									
QY	1	CGGACACCATGAAATGGGTGCAAAACCTTTGCGGGTATGGCATGATAGCGCCCGGAAGAGA	60						
DB	1	CGGACACCATGAAATGGGTGCAAAACCTTTGCGGGTATGGCATGATAGCGCCCGGAAGAGA	60						
QY	61	GTCAATTCAGGTGTGAATGTGAACCAAGTAAAGTATACGATGTGCGAGAGTATGCGG	120						
DB	61	GTCAATTCAGGTGTGAATGTGAACCAAGTAAAGTATACGATGTGCGAGAGTATGCGG	120						
QY	121	GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACCGCCAGCCACGCTTCTGCGAANA	180						
DB	121	GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACCGCCAGCCACGCTTCTGCGAANA	180						
QY	181	CGCGGAAAGTGAAGCGCGATGGCGAGCTGAATTTACATTTCCCAACCGCTGGCAC	240						
DB	181	CGCGGAAAGTGGNAGCGCGATGGCGAGCTGAATTTACATTTCCCAACCGCTGGCAC	240						
QY	241	AACAACTGGCGGCAAAACAGTCTGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300						
DB	241	AACAACTGGCGGCAAAACAGTCTGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300						
QY	301	ACGCGCGCTCGCAAAATTTGTCGGCGGATTAATCTCGGCCCATCAATGGGTGCCAGG	360						
DB	301	ACGCGCGCTCGCAAAATTTGTCGGCGGATTAATCTCGGCCCATCAATGGGTGCCAGG	360						
QY	361	TGTTGGTTCGATGGTAGAACGAGCGGCTGAAAGCCTGTAAAGCGGGGTGCACAATC	420						
DB	361	TGTTGGTTCGATGGTAGAACGAGCGGCTGAAAGCCTGTAAAGCGGGGTGCACAATC	420						
QY	421	TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTCCGCTGGATGACCAAGGATGCCA	480						
DB	421	TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTCCGCTGGATGACCAAGGATGCCA	480						
QY	481	TTGCTGTGGAGCTGCCTGCATTAATGTTCCGGCGTTATTTCTGATGTCCTGACCGA	540						
DB	481	TTGCTGTGGAGCTGCCTGCATTAATGTTTCGGCGGTTATTTCTGATGTCCTGACCGA	540						
QY	541	CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGGACTGGCGGTGGAGCATC	600						
DB	541	CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGGACTGGCGGTGGAGCATC	600						
QY	601	TGGTCGCATTTGGGTCAACGCAAAATCGCGCTGTAGCGGGCCCATTTAAGTTCCTGTCGG	660						
DB	601	TGGTCGCATTTGGGTCAACGCAAAATCGCGCTGTAGCGGGCCCATTTAAGTTCCTGTCGG	660						
QY	661	CGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
DB	661	CGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
QY	721	CGGAACGGGAAGGACGTGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA	780						
DB	721	CGGAACGGGAAGGACGTGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA	780						
QY	781	ATGAGGGCATCGTTCCTCCATCTGCGATGCTGTTTGCCTGCAACGATCAGATGGCGTGGCGCAA	840						
DB	781	ATGAGGGCATCGTTCCTCCATCTGCGATGCTGTTTGCCTGCAACGATCAGATGGCGTGGCGCAA	840						
QY	841	TGCGGCGCATACCGAGTCCGGCTGGCGTGGTGGGATATCTCGGTAGTGGGATAG	900						
DB	841	TGCGGCGCATACCGAGTCCGGCTGGCGTGGTGGGATATCTCGGTAGTGGGATAG	900						
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCATCAACAGGATTTTC	960						
DB	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCATCAACAGGATTTTC	960						
QY	961	GCCTGTGGGGCAAAACCGAGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGGGTGA	1020						
DB	961	GCCTGTGGGGCAAAACCGAGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGGGTGA	1020						
QY	1021	AGGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA	1080						

DB	1021	AGGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA	1080						
QY	1081	CGAAACCGCCCTCTCCCGCGCGTGGCCGATTCATTAAATGCAGTGGCACGACAGGTTT	1140						
DB	1081	CGAAACCGCCCTCTCCCGCGCGTGGCCGATTCATTAAATGCAGTGGCACGACAGGTTT	1140						
QY	1141	CCGACTGGAAGCGGCGAGTGAGGCAACGAATTAATGTGAGTTAGTCTCATATTAG	1200						
DB	1141	CCGACTGGAAGCGGCGAGTGAGGCAACGAATTAATGTGAGTTAGTCTCATATTAG	1200						
QY	1201	GCACAATTTCTATGTTTCAGACGCTTATCATGCAGTGCAGGTGCACCAATCTTCTGGCG	1260						
DB	1201	GCACAATTTCTATGTTTCAGACGCTTATCATGCAGTGCAGGTGCACCAATCTTCTGGCG	1260						
QY	1261	TCAGCAGCCATTCGGAAGCTGTGGTATGGCTGTGCAGGTCTGTAAATCACTGCATAAATTCG	1320						
DB	1261	TCAGCAGCCATTCGGAAGCTGTGGTATGGCTGTGCAGGTCTGTAAATCACTGCATAAATTCG	1320						
QY	1321	TGTGCTCAAGCGGCATCTCCGCTTCGGATTAATGTTTTTTCGGCCGCAATCAATACGGTT	1380						
DB	1321	TGTGCTCAAGCGGCATCTCCGCTTCGGATTAATGTTTTTTCGGCCGCAATCAATACGGTT	1380						
QY	1381	CTGSCAAATATTTCTGAAATGAGCTGTTCACAATTAATCATCGGCTCGTAAATGTGTGA	1440						
DB	1381	CTGSCAAATATTTCTGAAATGAGCTGTTCACAATTAATCATCGGCTCGTAAATGTGTGA	1440						
QY	1441	ATTGTGAGCGGATTAACAATTTTTCACAGGAAACGCCAGTCCGTTTAGTGTTCACGA	1500						
DB	1441	ATTGTGAGCGGATTAACAATTTTTCACAGGAAACGCCAGTCCGTTTAGTGTTCACGA	1500						
QY	1501	GCACTTTCACCAAGGACCATAGATTAATGAAATCGAAGAAAGGTAAATCTGGTAATCTGG	1560						
DB	1501	GCACTTTCACCAAGGACCATAGATTAATGAAATCGAAGAAAGGTAAATCTGGTAATCTGG	1560						
QY	1561	ATTAACGGCGATTAAGGCTATAACGGCTCTCGCTGAAGTCCGTAAGAAATTCAGAAAGAT	1620						
DB	1561	ATTAACGGCGATTAAGGCTATAACGGCTCTCGCTGAAGTCCGTAAGAAATTCAGAAAGAT	1620						
QY	1621	ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT	1680						
DB	1621	ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT	1680						
QY	1681	GCGCAACTGCGCATGCGCTTGACATTAATCTTCTGGGCAACGACCGCTTGGTGGCTAC	1740						
DB	1681	GCGCAACTGCGCATGCGCTTGACATTAATCTTCTGGGCAACGACCGCTTGGTGGCTAC	1740						
QY	1741	GCTCAATCTGGCTGTTGGCTGAATCAACCGGACCAAGCGTTCACGACCAAGCTGTAT	1800						
DB	1741	GCTCAATCTGGCTGTTGGCTGAATCAACCGGACCAAGCGTTCACGACCAAGCTGTAT	1800						
QY	1801	CCGTTTACCTGGGATGCGGTACGTTTCAACCGCAAGCTGATTGCTTACCCGATCGCTGTT	1860						
DB	1801	CCGTTTACCTGGGATGCGGTACGTTTCAACCGCAAGCTGATTGCTTACCCGATCGCTGTT	1860						
QY	1861	GAAGCGTTATCGCTGATTTATAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA	1920						
DB	1861	GAAGCGTTATCGCTGATTTATAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA	1920						
QY	1921	GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAGGTGAAGCGCGCTGATGTTCAAC	1980						
DB	1921	GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAGGTGAAGCGCGCTGATGTTCAAC	1980						
QY	1981	CTGCAAGAACCGTACTCTCACTGGCGCTGATTGCTGCTGAGCGGGGTTATCGCTTCAAG	2040						
DB	1981	CTGCAAGAACCGTACTCTCACTGGCGCTGATTGCTGCTGAGCGGGGTTATCGCTTCAAG	2040						
QY	2041	TATGAAAAACGCAAGTACGACATTAAGACGTTGGCGCTGGATTAACGCTGGCGCAAGCG	2100						
DB	2041	TATGAAAAACGCAAGTACGACATTAAGACGTTGGCGCTGGATTAACGCTGGCGCAAGCG	2100						
QY	2101	GGTCTGACCTTCTGGTTGACCTGATTTAAAAACAAACACATGAATGCGACGACCATTTAC	2160						

```
Db 2101 GGTCTGACCTTCTCGTGGTACCTGATTAAAAACAACACATGAATGACGACACCGGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCAACATCGACACACGCAAGAGTGAATTTATGGTGAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACACGCAAGAGTGAATTTATGGTGAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCAACACCGTTCCTTGGCGTGTGAGGCGAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGGTCAACCATCAACACCGTTCCTTGGCGTGTGAGGCGAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGTTTATAAGACAAACCGCTGGGTGCGGTGAGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGTTTATAAGACAAACCGCTGGGTGCGGTGAGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAGATCCACGATTGCGCCACCATGGNAACCGCCAGAAAGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGATTGCGCCACCATGGNAACCGCCAGAAAGTGAATCATG 2520
QY 2521 CCGAATCCCGCAGATGTCGCTTTCTGCTGATGCGGTGCTGCTGCTGATCAAGCC 2580
Db 2521 CCGAATCCCGCAGATGTCGCTTTCTGCTGATGCGGTGCTGCTGCTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGACACTGTGATGAAGCCCTGAAAGACGCGAGACTTAATCGAGTCTG 2640
Db 2581 GCCAGCGTCTGACACTGTGATGAAGCCCTGAAAGACGCGAGACTTAATCGAGTCTG 2640
QY 2641 AACAAACAACAATAAACAATAACAACAACCTCGGGATCGAGGGAAGGATTTTCAAGATTC 2700
Db 2641 AACAAACAACAATAAACAATAACAACAACCTCGGGATCGAGGGAAGGATTTTCAAGATTC 2700
QY 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709

RESULT 8
ADO23608
ID ADO23608 standard; DNA; 7259 BP.
AC ADO23608;
XX
DT 01-JUL-2004 (first entry)
XX
DE DNA encoding MBP-Toxop30del18(82-294aa) fusion protein.
XX
KW P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; gene.
XX
OS Toxoplasma gondii.
OS Synthetic.
XX
FN US2004067239-A1.
XX
PD 08-APR-2004.
XX
PF 02-OCT-2002; 2002US-00263153.
XX
PR 02-OCT-2002; 2002US-00263153.
XX
PA (MAIN/) MAINE G T.
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX
```

```
PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX WPI; 2004-304563/28.
DR P-PSDB; ADO23609.
XX
PT Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
PT antibodies to Toxoplasma gondii.
XX
XX Example 2; Fig 22; 114pp; English.
XX
CC The invention relates to a purified P30 antigen (I) chosen from 3 fully
CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
CC fusion proteins. (I) is useful for detecting the presence of IgM
CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC ToxoP30 fusion protein of the invention.
XX
XX Sequence 7259 BP; 1870 A; 1822 C; 1926 G; 1641 T; 0 U; 0 Other;
```

```
Query Match 81.5%; Score 2690.4; DB 12; Length 7259;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGCTGCAAAACCTTTTCGCGGTATGGCATGATAGCCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGCTGCAAAACCTTTTCGCGGTATGGCATGATAGCCCCGGAAGAGA 60
QY 61 GTCAAATTCAGGTTGGTGAATGTGAACACCAAGTATACGATGTGCGAGATATGCCG 120
Db 61 GTCAAATTCAGGTTGGTGAATGTGAACACCAAGTATACGATGTGCGAGATATGCCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCCGGTGTGTAACCAAGCGCCAGCCAGCTTTCTGCGAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCCGGTGTGTAACCAAGCGCCAGCCAGCTTTCTGCGAAA 180
QY 181 CGCGGGAAGAGTGGAGCGGCGATGCGGAGCTGAATTACATTCCTCCAAACCGGTGGCAC 240
Db 181 CGCGGGAAGAGTGGAGCGGCGATGCGGAGCTGAATTACATTCCTCCAAACCGGTGGCAC 240
QY 241 AACAACTGGCGGCAAAACAGTTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAAAACAGTTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTTCGCAAAATGTCGCGGCGATTAAATCTCGCCCGCATCAATCGGGTGCAGCG 360
Db 301 ACGCGCGTTCGCAAAATGTCGCGGCGATTAAATCTCGCCCGCATCAATCGGGTGCAGCG 360
QY 361 TGGTGTGTGTCGATGTTAGAACGAGCGGCTGGAAGCTGTAAAGCGGCGGTGCACAAATC 420
Db 361 TGGTGTGTGTCGATGTTAGAACGAGCGGCTGGAAGCTGTAAAGCGGCGGTGCACAAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATATATCCGCTGGATGACCAAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATATATCCGCTGGATGACCAAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTATTATTTCTGATGTCCTGACCAAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTATTATTTCTGATGTCCTGACCAAGA 540
QY 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGGGAGCATC 600
```

601 TGGTCGCAATTTGGGTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
601 TGGTCGCAATTTGGGTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
661 CGCGTCTCGCTGTGCTGTGCTGCGCAATAATATCTCTCACTCGCAATCAAAATTCAGCGCGATAG 720
661 CGCGTCTCGCTGTGCTGTGCTGCGCAATAATATCTCTCACTCGCAATCAAAATTCAGCGCGATAG 720
721 CGGAACGGGAAGGCGACCTGGAGTGCATGTCCGGTTTTCACAAACCATGCAAAATGCTGA 780
721 CGGAACGGGAAGGCGACCTGGAGTGCATGTCCGGTTTTCACAAACCATGCAAAATGCTGA 780
781 ATGAGGGCATGTTCCCACTGCGATGCTGTTGCGCAACGATGAGATGCGCTGGGCGCAA 840
781 ATGAGGGCATGTTCCCACTGCGATGCTGTTGCGCAACGATGAGATGCGCTGGGCGCAA 840
841 TCGCGGCCATTACCGAGTCCGGGCTCGCGTGTGGTGCAGATATCTCGGTAGTGGGATACG 900
841 TCGCGGCCATTACCGAGTCCGGGCTCGCGTGTGGTGCAGATATCTCGGTAGTGGGATACG 900
901 ACGATACCGAAGACAGCTCATGTTATATCCGCCGTTTAAACCAATCAAAACAGGATTTTC 960
901 ACGATACCGAAGACAGCTCATGTTATATCCGCCGTTTAAACCAATCAAAACAGGATTTTC 960
961 GCCTGCTGGGCAACACAGCTGGACCGCTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
961 GCCTGCTGGGCAACACAGCTGGACCGCTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACCAACCTGGGCGCCCAATA 1080
1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACCAACCTGGGCGCCCAATA 1080
1081 CGCAAAACCGCTCTCCCGCGGTTGGCCGATTCAATTAATGCACTGTCACACAGGTTT 1140
1081 CGCAAAACCGCTCTCCCGCGGTTGGCCGATTCAATTAATGCACTGTCACACAGGTTT 1140
1141 CCCGACTCGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGCTACTCATTTAG 1200
1141 CCCGACTCGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTTAGCTACTCATTTAG 1200
1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAAGTGCACCAATGCTCTGGCG 1260
1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAAGTGCACCAATGCTCTGGCG 1260
1261 TCAGGACGCCATCGGAAGCTGGTATGCGTGTGAGGTGCTGAGGTGCTGATTAATTCG 1320
1261 TCAGGACGCCATCGGAAGCTGGTATGCGTGTGAGGTGCTGAGGTGCTGATTAATTCG 1320
1321 TGTGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGCGACATCATACGGTT 1380
1321 TGTGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGCGACATCATACGGTT 1380
1381 CTGGCAAAATATTTGAAATGAGCTGTGACAATTAATCATCGGCTCGTATAATGTTGGA 1440
1381 CTGGCAAAATATTTGAAATGAGCTGTGACAATTAATCATCGGCTCGTATAATGTTGGA 1440
1441 ATTGTGAGCGGATAACAATTTACACAGGAACAGCCAGTCCGTTTAGTGTGTTTCAGCA 1500
1441 ATTGTGAGCGGATAACAATTTACACAGGAACAGCCAGTCCGTTTAGTGTGTTTCAGCA 1500
1501 GCACCTTACCAACAAAGGACATAGATTATGAAAACTGAAGAAAGGTAATCTGTAATCTGG 1560
1501 GCACCTTACCAACAAAGGACATAGATTATGAAAACTGAAGAAAGGTAATCTGTAATCTGG 1560
1561 ATTAACGCGATAAAGGCTATAACGCTCTCGCTGAAGTGGTAAGAAATTCAGAAAGAT 1620
1561 ATTAACGCGATAAAGGCTATAACGCTCTCGCTGAAGTGGTAAGAAATTCAGAAAGAT 1620
1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATATAAATTCGGAAGAGAAATTCACAGGTT 1680
1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATATAAATTCGGAAGAGAAATTCACAGGTT 1680
1681 GCGGCAACTGGCGATGGGCCCTGACATTAATCTCTGGGCAACAGCCGCTTTGGTGGCTAC 1740

RESULT 9
ADO23613
ID ADO23613 standard; DNA; 7322 BP.
XX AC ADO23613;

1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACAGCCGCTTTGGTGGCTAC 1740
1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCGGAAGAGCGTTTCAGGACAAGCTGTAT 1800
1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCGGAAGAGCGTTTCAGGACAAGCTGTAT 1800
1801 CCGTTTACCTGGGATGCGGTACGTTACAACGGCAAGCTGATTCCTTACCCGATCCTGTT 1860
1801 CCGTTTACCTGGGATGCGGTACGTTACAACGGCAAGCTGATTCCTTACCCGATCCTGTT 1860
1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGTAAGAGCGCTGATGTTCAAC 1980
1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGTAAGAGCGCTGATGTTCAAC 1980
1981 CTGCAAGAACCGTACTTACCTGGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
1981 CTGCAAGAACCGTACTTACCTGGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
2041 TATGAAAACGCAAGTACGACATTAAGAACGCTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
2041 TATGAAAACGCAAGTACGACATTAAGAACGCTGGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAAAACAACACATGAAATGSCAGACACCGATTAC 2160
2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAAAACAACACATGAAATGSCAGACACCGATTAC 2160
2161 TCCATCGAAGAGTGCCTTTAATAAAGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
2161 TCCATCGAAGAGTGCCTTTAATAAAGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
2221 GCATGGTCCAAACATCGACACAGCAAGTGAATTAATGTTGTAACCGGTACTGCGGACCTTC 2280
2221 GCATGGTCCAAACATCGACACAGCAAGTGAATTAATGTTGTAACCGGTACTGCGGACCTTC 2280
2281 AAGGTCACCAACCATCCAAACCGTTTCTGGCGTGTCTGAGCGAGGTATTAAACGCCCCAGT 2340
2281 AAGGTCACCAACCATCCAAACCGTTTCTGGCGTGTCTGAGCGAGGTATTAAACGCCCCAGT 2340
2341 CCGAACAAGAGTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCGTG 2400
2341 CCGAACAAGAGTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCGTG 2400
2401 GAAGCGTTAATAAGACAAACCCCTGGGTGCGGTAGCGCTGAAAGTCTTTACGAGGAAG 2460
2401 GAAGCGTTAATAAGACAAACCCCTGGGTGCGGTAGCGCTGAAAGTCTTTACGAGGAAG 2460
2461 TTGGCAAAAGATCCACGTTATTTGCCGCCACCATGGAACCAACGCCAGAAAGGTGAAATCATG 2520
2461 TTGGCAAAAGATCCACGTTATTTGCCGCCACCATGGAACCAACGCCAGAAAGGTGAAATCATG 2520
2521 CCGAACAATCCCGCAGATGTCGCGCTTTCTGGTATGCCGTGCTGCTGATCAACGCGC 2580
2521 CCGAACAATCCCGCAGATGTCGCGCTTTCTGGTATGCCGTGCTGCTGATCAACGCGC 2580
2581 GCCAGCGTCTGACAGCTGTCGATGAGCCCTGGAAGACGCGCAGACTTAATTCGAGCTCG 2640
2581 GCCAGCGTCTGACAGCTGTCGATGAGCCCTGGAAGACGCGCAGACTTAATTCGAGCTCG 2640
2641 AACCAACAACAATAACAAATAACAAACCTCGGGATCGAGGGAGGATTTTCAGAAATTC 2700
2641 AACCAACAACAATAACAAATAACAAACCTCGGGATCGAGGGAGGATTTTCAGAAATTC 2700

XX 01-JUL-2004 (first entry)
 XX DNA encoding MBP-Toxop30del10(52-284aa) fusion protein.
 XX DE
 XX KW P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
 KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
 KW Toxoplasma gondii; ds; gene.
 XX OS
 OS Toxoplasma gondii.
 OS Synthetic.
 XX US2004067239-A1.
 XX PN
 XX PD 08-APR-2004.
 XX XX
 XX PF 02-OCT-2002; 2002US-00263153.
 XX XX
 XX PR 02-OCT-2002; 2002US-00263153.
 XX XX
 XX PA (MAIN/) MAINE G T.
 PA (PATE/) PATEL C B.
 PA (GINS/) GINSBURG S R.
 PA (BLIE/) BLIESE T R.
 XX
 PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
 XX WPI; 2004-304563/28.
 DR P-PSDB; ADO23614.
 XX
 XX Novel purified polypeptide having sequence identity to amino acid
 PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
 PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
 PT antibodies to Toxoplasma gondii.
 PS Example 2; Fig 25; 114pp; English.
 XX
 CC The invention relates to a purified P30 antigen (I) chosen from 3 fully
 CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
 CC amino acids added to the C-terminus of the amino acid sequence of
 CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
 CC which at least one of the five C-terminal cysteine amino acids of the
 CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
 CC with alanine, or comprising the amino acid sequence chosen from MBP-
 CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
 CC fusion proteins. (I) is useful for detecting the presence of IgM
 CC antibodies to Toxoplasma gondii in a test sample, which involves
 CC contacting the test sample suspected of containing the IgM antibodies
 CC with a composition comprising (I) and detecting the presence of (I)/IgM
 CC antibody complexes. The present sequence represents DNA encoding a MBP-
 CC ToxoP30 fusion protein of the invention.
 XX
 SQ Sequence 7322 BP; 1887 A; 1844 C; 1934 G; 1657 T; 0 U; 0 Other;
 Query Match 81.5%; Score 2690.4; DB 12; Length 7322;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CCGACACCATCGAATGGTGAACAAACCTTTCCCGGTATGGCATGATAGCGCCCGGAAGAGA 60
 DB
 QY 1 CCGACACCATCGAATGGTGAACAAACCTTTCCCGGTATGGCATGATAGCGCCCGGAAGAGA 60
 DB
 QY 61 GTCAATTCAGGGTGGTGAATGTGAACACAGTAAAGTTATACGATGTCGAGAGATGCGG 120
 DB 61 GTCAATTCAGGGTGGTGAATGTGAACACAGTAAAGTTATACGATGTCGAGAGATGCGG 120
 QY 121 GTGTCTCTTATCAGACCGGTTTCCCGCGTGTGAACACAGTAAAGTTATACGATGTCGAGAG 180
 DB 121 GTGTCTCTTATCAGACCGGTTTCCCGCGTGTGAACACAGTAAAGTTATACGATGTCGAGAG 180
 QY 181 CGCGGGGAAAAGTGAACCGCGGATGCGGAGTGAATTCATTCGCAACCGCGTGGCAC 240
 DB 181 CGCGGGGAAAAGTGAACCGCGGATGCGGAGTGAATTCATTCGCAACCGCGTGGCAC 240

QY 241 AACAACTGGCGGGCAAAACAGTCGTTGTGATTTGGCGTTGCACCTCCAGTCTGCGCCCTGC 300
 DB 241 AACAACTGGCGGGCAAAACAGTCGTTGTGATTTGGCGTTGCACCTCCAGTCTGCGCCCTGC 300
 QY 301 ACGCGCGTTCGCAAAATTTGTCGCGCGATTAATCTTCGCGCGGATCAACTGGGTGCGACGC 360
 DB 301 ACGCGCGTTCGCAAAATTTGTCGCGCGATTAATCTTCGCGCGGATCAACTGGGTGCGACGC 360
 QY 361 TGGTGGTGTGATGTTAGAACGAGCGCGTTCGAAGCGCTGTAAGCGCGCGTGCACAATC 420
 DB 361 TGGTGGTGTGATGTTAGAACGAGCGCGTTCGAAGCGCTGTAAGCGCGCGTGCACAATC 420
 QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAACATATCCGCTGGATGACAGGATGCCA 480
 DB 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAACATATCCGCTGGATGACAGGATGCCA 480
 QY 481 TTGCTGTGGAAGCTGCTGCACATAATGTTCCGGGTTATTTCTTGATGTCTCTGACCAGA 540
 DB 481 TTGCTGTGGAAGCTGCTGCACATAATGTTCCGGGTTATTTCTTGATGTCTCTGACCAGA 540
 QY 541 CACCATCAACAGTATTAATTTCTCCCATGAAGACGCTACGCGACTGGGCGTGGAGCATC 600
 DB 541 CACCATCAACAGTATTAATTTCTCCCATGAAGACGCTACGCGACTGGGCGTGGAGCATC 600
 QY 601 TGGTTCGATTTGGGTTCACAGCAATCGCGTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
 DB 601 TGGTTCGATTTGGGTTCACAGCAATCGCGTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
 QY 661 CGGCTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCCATAG 720
 DB 661 CGGCTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCCATAG 720
 QY 721 CGGAACGGGAAGCGGACGTGAGTGCCATGTCGGTTTTCAACAAACCATGCATGCTGA 780
 DB 721 CGGAACGGGAAGCGGACGTGAGTGCCATGTCGGTTTTCAACAAACCATGCATGCTGA 780
 QY 781 ATGAGGCGATCGTTCCCACTCGCATCTGTTGCCAACGATCAGATGGCGCTGGCGGCAA 840
 DB 781 ATGAGGCGATCGTTCCCACTCGCATCTGTTGCCAACGATCAGATGGCGCTGGCGGCAA 840
 QY 841 TGGCGCGCATTAACGAGTCCGGCTCGCGTGTGGTCCGGATATCTCGTAGTGGGATACG 900
 DB 841 TGGCGCGCATTAACGAGTCCGGCTCGCGTGTGGTCCGGATATCTCGTAGTGGGATACG 900
 QY 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAACCCACCATCAACAGGATTTTC 960
 DB 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAACCCACCATCAACAGGATTTTC 960
 QY 961 GCCTGCTGGGGCAAAACGAGTGGACCGCTGTGCAACTCTCTCAGGGCCAGGGCGTGA 1020
 DB 961 GCCTGCTGGGGCAAAACGAGTGGACCGCTGTGCAACTCTCTCAGGGCCAGGGCGTGA 1020
 QY 1021 AGGGCAATCAGCTGTGTCGGCTCTCACTGGTGAAGAAAAACCACTGGCGCCCAATA 1080
 DB 1021 AGGGCAATCAGCTGTGTCGGCTCTCACTGGTGAAGAAAAACCACTGGCGCCCAATA 1080
 QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTTCAATTAATGAGTGGCGACAGCAGGTTT 1140
 DB 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTTCAATTAATGAGTGGCGACAGCAGGTTT 1140
 QY 1141 CCGCATGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCTCACTCATTAG 1200
 DB 1141 CCGCATGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCTCACTCATTAG 1200
 QY 1201 GCACAAATTCATGTTTGAACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260
 DB 1201 GCACAAATTCATGTTTGAACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260
 QY 1261 TCAGGCGAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCGAGTCTGTAATCACTGATATTCG 1320
 DB 1261 TCAGGCGAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCGAGTCTGTAATCACTGATATTCG 1320

QY 1321 TGTCTCAAGCGCAGCTCCCGTTCTGGATAATGTTTTTGGCCGACATCATACGGTT 1380
Db 1321 TGTCTCAAGCGCAGCTCCCGTTCTGGATAATGTTTTTGGCCGACATCATACGGTT 1380
QY 1381 CTGGCAATATTCTGAAATGAGCTGTTGACAATTAATCATCGGCTCGTATAATGTTGGA 1440
Db 1381 CTGGCAATATTCTGAAATGAGCTGTTGACAATTAATCATCGGCTCGTATAATGTTGGA 1440
QY 1441 ATTGTAGCGGATAACAATTTTACACAGGAAACAGCCAGTCGGTTTAGTGTTTTTCAAGA 1500
Db 1441 ATTGTAGCGGATAACAATTTTACACAGGAAACAGCCAGTCGGTTTAGTGTTTTTCAAGA 1500
QY 1501 GCATCTCACCAACAGGACCATAGATTATGAAAACTGAAGAGTTAAACTGTAATCTGG 1560
Db 1501 GCATCTCACCAACAGGACCATAGATTATGAAAACTGAAGAGTTAAACTGTAATCTGG 1560
QY 1561 ATTTAAGCGGATAAAGGCTATAACGGTCTCGCTGAAAGTCGGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTTAAGCGGATAAAGGCTATAACGGTCTCGCTGAAAGTCGGTAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCGGAAGAGAAATTCCTCCAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCGGAAGAGAAATTCCTCCAGGTT 1680
QY 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACAGCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACAGCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGTGCTCAATCAACCCGGACAAAGCGTTCCAGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGTGCTCAATCAACCCGGACAAAGCGTTCCAGACAAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTACACGCAAGCTGATGCTTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACACGCAAGCTGATGCTTTACCCGATCGCTGT 1860
QY 1861 GAAGGTTATCGCTGATTTATAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGGTTATCGCTGATTTATAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCTGTGATAAAGACTGAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCTGTGATAAAGACTGAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGACGCGGGTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGACGCGGGTTATCGGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGCGTGGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGCGTGGGATTAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACATCAATGACACACCGATAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACATCAATGACACACCGATAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGCGCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGCGCGCTGG 2220
QY 2221 GCATGTCACATCGACACAGCAAGTGAATTAATGTTGTTACGGTACTCGCGACCTTC 2280
Db 2221 GCATGTCACATCGACACAGCAAGTGAATTAATGTTGTTACGGTACTCGCGACCTTC 2280
QY 2281 AAGGTCACCATCAACCGTTGCTGGCGTCTGAGCGAGGTTAATTAACCGCCAGT 2340
Db 2281 AAGGTCACCATCAACCGTTGCTGGCGTCTGAGCGAGGTTAATTAACCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460

Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGTGAAGTCTTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCACGCTATTGCGCCCAACCATGGAAGAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGCTATTGCGCCCACTATGGAAGAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGGTCGCTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGGTCGCTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGGTCGTCAGACTGTCGATGAAGCCCTGGAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTCGTCAGACTGTCGATGAAGCCCTGGAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACACAAACAATAAACAATAACAACACCTCGGGATCGAGGGAGGATTTTCAGAAATTC 2700
Db 2641 AACACAAACAATAAACAATAACAACACCTCGGGATCGAGGGAGGATTTTCAGAAATTC 2700

RESULT 10
ADO23603
ID ADO23603 standard; DNA; 7352 BP.
XX
AC ADO23603;
XX
DT 01-JUL-2004 (first entry)
XX
DE DNA encoding MBP-ToxoP30del14C(52-294aa) fusion protein.
KW P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; gene.
OS Toxoplasma gondii.
OS Synthetic.
XX
PN US2004067239-A1.
XX
PD 08-APR-2004.
XX
PF 02-OCT-2002; 2002US-00263153.
XX
PR 02-OCT-2002; 2002US-00263153.
XX
PA (MAIN/) MAINE G T.
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX
PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX
DR WPI: 2004-304563/28.
DR P-PSDB; ADO23604.
XX
PT Novel purified polypeptide having sequence identity to amino acid sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C, Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG antibodies to Toxoplasma gondii.
XX
PS Example 2; Fig 19; 114pp; English.
XX
CC The invention relates to a purified P30 antigen (I) chosen from 3 fully defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6 amino acids added to the C-terminus of the amino acid sequence of Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in which at least one of the five C-terminal cysteine amino acids of the amino acid sequence of Toxo30del13C P30 antigen sequence is substituted with alanine, or comprising the amino acid sequence chosen from MBP-Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1 fusion proteins. (I) is useful for detecting the presence of IgM antibodies to Toxoplasma gondii in a test sample, which involves contacting the test sample suspected of containing the IgM antibodies

CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
XX Toxop30 fusion protein of the invention.

SQ Sequence 7352 BP; 1896 A; 1852 C; 1943 G; 1661 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7352;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY	1	CGGACACCATCAATGGTGCAGAACTTTCCGGGTATGCGATGATAGGCCCGGAAGAGA	60
DB	1	CGGACACCATCAATGGTGCAGAACTTTCCGGGTATGCGATGATAGGCCCGGAAGAGA	60
QY	61	GTCAATTCAGGTTGAGTAAAGTGAACACAGTAACGTTATACGATGTCGACAGTATGCCG	120
DB	61	GTCAATTCAGGTTGAGTAAAGTGAACACAGTAACGTTATACGATGTCGACAGTATGCCG	120
QY	121	GGTCTCTTATCAGACCGTTTCCCGGTGGTGAACAGGCCAGCCAGCGTTTCTGCGAAAA	180
DB	121	GGTCTCTTATCAGACCGTTTCCCGGTGGTGAACAGGCCAGCCAGCGTTTCTGCGAAAA	180
QY	181	CGCGGAAAAAGTGAAGCGGATGGCGAGCTGAATTTACATTTCCCAACCGCTGGCAC	240
DB	181	CGCGGAAAAAGTGAAGCGGATGGCGAGCTGAATTTACATTTCCCAACCGCTGGCAC	240
QY	241	AACAACTGCGGGCAAAAGTGTGCTGATTTGGGTTGCCACTCCAGTCTGGCCCTGC	300
DB	241	AACAACTGCGGGCAAAAGTGTGCTGATTTGGGTTGCCACTCCAGTCTGGCCCTGC	300
QY	301	ACGCCCGTCGCAAAATTCGCGGGATTAATCTCGCGCGATCACTGGGTGCGACGC	360
DB	301	ACGCCCGTCGCAAAATTCGCGGGATTAATCTCGCGCGATCACTGGGTGCGACGC	360
QY	361	TGGTGGTTCGATGTAGAACGAAGCGCGTGAAGCGCTGAAGCGCGGTGCACAAATC	420
DB	361	TGGTGGTTCGATGTAGAACGAAGCGCGTGAAGCGCTGAAGCGCGGTGCACAAATC	420
QY	421	TTCTCGCCCAAGCGTCACTAGTGGCTGATCAATTAATCTCCGCTGGATGACCAAGATGCCA	480
DB	421	TTCTCGCCCAAGCGTCACTAGTGGCTGATCAATTAATCTCCGCTGGATGACCAAGATGCCA	480
QY	481	TTGCTGTGGAAGCTCCCTGCACTAATGTTCCGGGTATTTCTTGATGCTCTGACCCAGA	540
DB	481	TTGCTGTGGAAGCTCCCTGCACTAATGTTCCGGGTATTTCTTGATGCTCTGACCCAGA	540
QY	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGCACTGGCGGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGCACTGGCGGTGGAGCATC	600
QY	601	TGGTCCGATTTGGTCAACGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTCTGCTCGG	660
DB	601	TGGTCCGATTTGGTCAACGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTCTGCTCGG	660
QY	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAAACGGGAAGCGCATGGAAGTCCCATGTCGGTTTTTCAACAAACCATGCAATGCTGA	780
DB	721	CGGAAACGGGAAGCGCATGGAAGTCCCATGTCGGTTTTTCAACAAACCATGCAATGCTGA	780
QY	781	ATGAGGCGATCTTCCCACTCGCATGCTGGTTCGAACGATCAGATGGCGCTGGCGCAA	840
DB	781	ATGAGGCGATCTTCCCACTCGCATGCTGGTTCGAACGATCAGATGGCGCTGGCGCAA	840
QY	841	TGCGGCCATTTACCGAGTCCGGCTGCGGTGGTGGCGGATATCTCGGTAGTGGATACG	900
DB	841	TGCGGCCATTTACCGAGTCCGGCTGCGGTGGTGGCGGATATCTCGGTAGTGGATACG	900
QY	901	ACGATACCGAGACAGTCTATGTTATATCCGCGGTTAACCAATCAACAGGATTTTC	960
DB	901	ACGATACCGAGACAGTCTATGTTATATCCGCGGTTAACCAATCAACAGGATTTTC	960

QY	961	GCCTGCTGGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCCACGCGGTGA	1020
DB	961	GCCTGCTGGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCCACGCGGTGA	1020
QY	1021	AGGCAATCAGCTGTGCCCCGTCTCACTGGTGAAGAAAAACACCCCTGGCGCCCAATA	1080
DB	1021	AGGCAATCAGCTGTGCCCCGTCTCACTGGTGAAGAAAAACACCCCTGGCGCCCAATA	1080
QY	1081	CGAAAAACCGCTCTCCCGCGCGTTGGCGGATCATTAATGACGTGGCACAGAGTTT	1140
DB	1081	CGAAAAACCGCTCTCCCGCGCGTTGGCGGATCATTAATGACGTGGCACAGAGTTT	1140
QY	1141	CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGATTAAGTCACTCATTA	1200
DB	1141	CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGATTAAGTCACTCATTA	1200
QY	1201	GCACAAATCTCATGTTTGGACAGCTTATCATCGACTGCAACGCTGCAACATGCTTCTGG	1260
DB	1201	GCACAAATCTCATGTTTGGACAGCTTATCATCGACTGCAACGCTGCAACATGCTTCTGG	1260
QY	1261	TCAGGACGCAATCGAAAGCTGTGGTATGGCTGTGAGGTGCTAAATCACTGCAATAATTC	1320
DB	1261	TCAGGACGCAATCGAAAGCTGTGGTATGGCTGTGAGGTGCTAAATCACTGCAATAATTC	1320
QY	1321	TGTCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCAACCGTT	1380
DB	1321	TGTCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCAACCGTT	1380
QY	1381	CTGGCAAAATCTGAAATGAGCTGTGTGAACAATTAATCATCGGCTGCTATAATGTGTGA	1440
DB	1381	CTGGCAAAATCTGAAATGAGCTGTGTGAACAATTAATCATCGGCTGCTATAATGTGTGA	1440
QY	1441	ATTGTGAGCGGATTAACAATTTTCAACAGGAAACAGCCAGTCCGTTTATAGTGTTCACGA	1500
DB	1441	ATTGTGAGCGGATTAACAATTTTCAACAGGAAACAGCCAGTCCGTTTATAGTGTTCACGA	1500
QY	1501	GCATTTCCACCAACAAAGGACCATAGATTATGAAACTGAAGAAGTAAACTGGTAAATCTGG	1560
DB	1501	GCATTTCCACCAACAAAGGACCATAGATTATGAAACTGAAGAAGTAAACTGGTAAATCTGG	1560
QY	1561	ATTAAACGGCGATTAAGGCTATAACCGTCTCGTGAAGTCCGTTAAGAAATTCGAGAAAGAT	1620
DB	1561	ATTAAACGGCGATTAAGGCTATAACCGTCTCGTGAAGTCCGTTAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGTAAACTCGAAGAGAAATTCACACAGTT	1680
DB	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGTAAACTCGAAGAGAAATTCACACAGTT	1680
QY	1681	GGCGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCAACAGACCGCTTTGGTGGCTAC	1740
DB	1681	GGCGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCAACAGACCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCCCTGTGGCTGAATACCCCGGACAAAGCGTTCCAGAACAGCTGTAT	1800
DB	1741	GCTCAATCTGGCCCTGTGGCTGAATACCCCGGACAAAGCGTTCCAGAACAGCTGTAT	1800
QY	1801	CCGTTTACCTGGGATCCCGTACGTTACAAAGCAAGCTGATTTACCGGATCCGCTGTT	1860
DB	1801	CCGTTTACCTGGGATCCCGTACGTTACAAAGCAAGCTGATTTACCGGATCCGCTGTT	1860
QY	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTCCGAAACCCGCAAAAAACCTGGAA	1920
DB	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTCCGAAACCCGCAAAAAACCTGGAA	1920
QY	1921	GAGATCCCGCGCTGGATAAGAACTGAAGCGAAGGTAAAGCGCGCTGATGTTCAAC	1980
DB	1921	GAGATCCCGCGCTGGATAAGAACTGAAGCGAAGGTAAAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTTCCCTGGCGCTGATTTGCTGCTGACGGGGTTATCGGTTCAAG	2040
DB	1981	CTGCAAGAACCGTACTTCACTTCCCTGGCGCTGATTTGCTGCTGACGGGGTTATCGGTTCAAG	2040

Db 541 CACCCATCAACAGTATTATTTCTCCATGAAGACGGTACGCGACTGGCGCTGGAGCATC 600
Qy 601 TGGTCGCAATGGGTACACGAAATCGCGCTGTGTAGCGGCCCATTTAAAGTTCGTGCTCGG 660
Db 601 TGGTCGCAATGGGTACACGAAATCGCGCTGTGTAGCGGCCCATTTAAAGTTCGTGCTCGG 660
Qy 661 CGCGTCTGGCTGTGGCTGGCTGGCTGAATAATATCTCACTCGCAATCAAATTCAGCCGATAG 720
Db 661 CGCGTCTGGCTGTGGCTGGCTGGCTGAATAATATCTCACTCGCAATCAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGCGCACTCGAGTGCATGCTCGGTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGCGCACTCGAGTGCATGCTCGGTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATCTGTTCCCACTGGGATGTGTTGCCAAACGATCAGATGGCGCTGGCGCGAA 840
Db 781 ATGAGGGCATCTGTTCCCACTGGGATGTGTTGCCAAACGATCAGATGGCGCTGGCGCGAA 840
Qy 841 TGGCGGCCATTAACGAGTCCGGCTCGCGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGGCCATTAACGAGTCCGGCTCGCGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
Qy 961 GCCTGCTGGGCAAAACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGCTGGGCAAAACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGCGCGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGCGCGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGGCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGCAGCTGSCACGACAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGCAGCTGSCACGACAGTTT 1140
Qy 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Qy 1201 GCACAATTCATGTTTCACAGCTTATCATGACGTGCACGCTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTCATGTTTCACAGCTTATCATGACGTGCACGCTGCACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTAAATCACTGCATAATTTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTAAATCACTGCATAATTTCG 1320
Qy 1321 TGTGCTCAAGCGCATCCCGCTTCTGGATTAATGTTTTCGCGCGAGCATCAACCGTT 1380
Db 1321 TGTGCTCAAGCGCATCCCGCTTCTGGATTAATGTTTTCGCGCGAGCATCAACCGTT 1380
Qy 1381 CTGGCAATATTTCTGAATGAGCTGTTGACAATTAATCATCGGCTCGTAAATGTTGTTGGA 1440
Db 1381 CTGGCAATATTTCTGAATGAGCTGTTGACAATTAATCATCGGCTCGTAAATGTTGTTGGA 1440
Qy 1441 ATTGTGAGCGGATTAACAAATTTCAACAGGAACAGCCAGTCCGTTTAGTGTGTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATTAACAAATTTCAACAGGAACAGCCAGTCCGTTTAGTGTGTTTTCACGA 1500
Qy 1501 GCATTCACCAACAGGACCATAGATTATGAAACTGAAGATTAACCTGGTAAATCTGG 1560
Db 1501 GCATTCACCAACAGGACCATAGATTATGAAACTGAAGATTAACCTGGTAAATCTGG 1560
Qy 1561 ATTAACCGCGATTAAGGCTATAACGCTCTCGCTGAACTCGGTGAAGAAATTCGAGAAGAT 1620
Db 1561 ATTAACCGCGATTAAGGCTATAACGCTCTCGCTGAACTCGGTGAAGAAATTCGAGAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCACACAGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCACACAGTT 1680

RESULT 12

AD023639

ID AD023639 standard; DNA; 7370 BP.

Qy 1681 GCGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCCGCAAAAGGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCCGCAAAAGGTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTTACAACGGCAAGCTGATTTGCTTACC CGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACAACGGCAAGCTGATTTGCTTACC CGATCGCTGTT 1860
Qy 1861 GAAGGCTTATCGCTGATTTATACAAAGATCTGCTGCGAACCCCGCAAAACCTGGGAA 1920
Db 1861 GAAGGCTTATCGCTGATTTATACAAAGATCTGCTGCGAACCCCGCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGCGCCCTGATTTGCTGCTGACCGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCCCTGATTTGCTGCTGACCGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAAACGGCAAGTACGACATTAAGACCTGAAAGCGCGCTGATTAACCGCGGAAACG 2100
Db 2041 TATGAAAAACGGCAAGTACGACATTAAGACCTGAAAGCGCGCTGATTAACCGCGGAAACG 2100
Qy 2101 GGTCTGACCTTCTCTGTTGACCTGATTTAAATAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTTAAATAACAAACACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAAGAGTGCCTTTTAAAGCGGAAACAGCGATGACCATCAACCGCCGCTGG 2220
Db 2161 TCCATCGCAAGAGTGCCTTTTAAAGCGGAAACAGCGATGACCATCAACCGCCGCTGG 2220
Qy 2221 GCATGTTCCAAACATCGACACAGCAAGTGAATTTATGTTGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGTTCCAAACATCGACACAGCAAGTGAATTTATGTTGTAACGGTACTGCGGACCTTC 2280
Qy 2281 AAGGTCACCATCCAAAACCGTTTGGCGTCTGAGCGCAGGTATTAACCGCCGCTAG 2340
Db 2281 AAGGTCACCATCCAAAACCGTTTGGCGTCTGAGCGCAGGTATTAACCGCCGCTAG 2340
Qy 2341 CCCAAACAAAGAGCTGCAAAAGAGTTCCCTGCAAACTATCTGCTGATGTAAGGCTCG 2400
Db 2341 CCCAAACAAAGAGCTGCAAAAGAGTTCCCTGCAAACTATCTGCTGATGTAAGGCTCG 2400
Qy 2401 GAAGCGTTTAAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGTTTAAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGGCGAAAGATCCAGTATTCGCCGCCACATGGAAGAAACGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTCGCCGCCACATGGAAGAAACGCCAGAAAGGTGAAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGCTATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Qy 2581 GCCAGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640
Db 2581 GCCAGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640
Qy 2641 AACCAACAAACAAATAACCAATTAACCAACCTCGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACCAACAAACAAATAACCAATTAACCAACCTCGGATCGAGGGAAGGATTTTCAGAAATTC 2700

```
XX AC ADO23639;
XX AC
XX DT 01-JUL-2004 (first entry)
XX DE DNA encoding MBP-ToxoP30MIX1 fusion protein.
XX KW P30 antigen; Toxo30del13C; Toxo30del2C; ToxoP30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; db; gene.
XX OS Toxoplasma gondii.
XX OS Synthetic.
XX PW US2004067239-A1.
XX PD 08-APR-2004.
XX PF 02-OCT-2002; 2002US-00263153.
XX PR 02-OCT-2002; 2002US-00263153.
XX PA (MAIN/) MAINE G T.
XX PA (PATE/) PATEL C B.
XX PA (GINS/) GINSBURG S R.
XX PA (BLIE/) BLIESE T R.
XX FI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX WI WPI; 2004-304563/28.
XX DR P-PSDB; ADO23640.
XX PT Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
PT antibodies to Toxoplasma gondii.
XX Example 5; Fig 31; 114pp; English.
XX The invention relates to a purified P30 antigen (I) chosen from 3 fully
CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
CC fusion proteins. (I) is useful for detecting the presence of IgM
CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC ToxoP30 fusion protein of the invention.
XX Sequence 7370 BP; 1900 A; 1856 C; 1950 G; 1664 T; 0 U; 0 Other;
SQ
Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTCGGGGTATGCGCATGATAGCGCCCGGAAGAGA 60
DB 1 CCGACACCATCGAATGGTGCAGAAACCTTTCGGGGTATGCGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTACAGGTGTGAATGTGAACAGTAACTATACGATGTCGCGAGATATGCG 120
DB 61 GTCAATTACAGGTGTGAATGTGAACAGTAACTATACGATGTCGCGAGATATGCG 120
QY 121 GTGTCTTATCAGACCGGTTTCCCGGGTGGTGAACAGGCCAGCCACGTTTCTGCGAAAA 180
DB 121 GTGTCTTATCAGACCGGTTTCCCGGGTGGTGAACAGGCCAGCCACGTTTCTGCGAAAA 180
QY 181 CCGCGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTACATTCCTCCAAACCGGTGGCAC 240
```

```
Db 181 CCGCGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTACATTCCTCCAAACCGGTGGCAC 240
QY 241 AACAACTGGCGGCAAAACAGTCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAAAACAGTCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTTCGCAAAATTTGTCGGGGGATTAATCTCGCGCGGATCAATCGGTGTCGAGG 360
Db 301 ACGCGCGTTCGCAAAATTTGTCGGGGGATTAATCTCGCGCGGATCAATCGGTGTCGAGG 360
QY 361 TGGTGGTGTGATGTAAGAAAGAGCGCGTTCGAAGCTGTAAAGCGCGGTGCACAATC 420
Db 361 TGGTGGTGTGATGTAAGAAAGAGCGCGTTCGAAGCTGTAAAGCGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACGCGTCACTGAGTGGGCTGATCATTAACATATCCGCTGGATGACCAAGTCCA 480
Db 421 TTCTCGCGCAACGCGTCACTGAGTGGGCTGATCATTAACATATCCGCTGGATGACCAAGTCCA 480
QY 481 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTTCTGACCGA 540
Db 481 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTTCTGACCGA 540
QY 541 CACCCATCAACAGTATTAATTTTCTCCATGAAGAGCGTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTTCTCCATGAAGAGCGTACGCGACTGGGCGTGGAGCATC 600
QY 601 TGGTGGCATTTGGGTCAACAGCAAAATTCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCGG 660
Db 601 TGGTGGCATTTGGGTCAACAGCAAAATTCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCGG 660
QY 661 CGCGTCTCGCTGCTGGCTGGGCAATAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db 661 CGCGTCTCGCTGCTGGCTGGGCAATAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
QY 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATGCTTCCCACTCGCATGCTGGTTGCCAACGATCAGATGCGCTGGCGGCA 840
Db 781 ATGAGGGCATGCTTCCCACTCGCATGCTGGTTGCCAACGATCAGATGCGCTGGCGGCA 840
QY 841 TGGCGGCCATTAACGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
Db 841 TGGCGGCCATTAACGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCCACCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCCACCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAAACCACTTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAAACCACTTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCTGCGACAGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCTGCGACAGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAATTTCTATGTTTGAAGCTTATCATCACTGCAAGTGCAGGTGCAACCAATGTTCTGGG 1260
Db 1201 GCACAATTTCTATGTTTGAAGCTTATCATCACTGCAAGTGCAGGTGCAACCAATGTTCTGGG 1260
QY 1261 TCAGGCAGCCATCGGAAGCTGTGTTATGCTGTGCGAGTGTGAATCACTGATCAATTCG 1320
```

Db 1261 TCAGGAGCCATCGGAAGCTGTGGTATAGCTGTGCAGGTCGTAATCACTGCATAATTCCG 1320
Qy 1321 TGTGCTCAAGGCGCACTCCGCTTCGGGATAATGTTTTTTCGCCGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGGCGCACTCCGCTTCGGGATAATGTTTTTTCGCCGACATCATACGGTT 1380
Qy 1381 CTGGCAATATTTCTGAATAGCTGTGTGAATTAATCATCGGCTCGTATAATGTGTGA 1440
Db 1381 CTGGCAATATTTCTGAATAGCTGTGTGAATTAATCATCGGCTCGTATAATGTGTGA 1440
Qy 1441 ATTGTAGCGGATACAAATTTTCCACAGGAACGCGAGTCGGTTAGTGTTTTTCACGA 1500
Db 1441 ATTGTAGCGGATACAAATTTTCCACAGGAACGCGAGTCGGTTAGTGTTTTTCACGA 1500
Qy 1501 GCATTTCCAAACAAGGACATAGATTAATGAAGAACTGAAGAGTAAACTGTGTAATCGG 1560
Db 1501 GCATTTCCAAACAAGGACATAGATTAATGAAGAACTGAAGAGTAAACTGTGTAATCGG 1560
Qy 1561 ATTAACGGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCCAGCTTGAGCATCCGATTAACCTGGAAGAGAAATTCGCAAGGTT 1680
Db 1621 ACCGGAATTAAGTCCAGCTTGAGCATCCGATTAACCTGGAAGAGAAATTCGCAAGGTT 1680
Qy 1681 GCGGCACTGGCGATGGCCCTGACATTAATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCACTGGCGATGGCCCTGACATTAATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCTGTGGCTGAATATCAACCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTGGCTGAATATCAACCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACTCGGATGCGGTACGTTTACAAACGCAAGCTGATGTTGTTACCGATCGCTGTT 1860
Db 1801 CCGTTTACTCGGATGCGGTACGTTTACAAACGCAAGCTGATGTTGTTACCGATCGCTGTT 1860
Qy 1861 GAAGGTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCGCCCAAAACCTGGGAA 1920
Db 1861 GAAGGTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCGCCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCGGCGCTGGATAAGAACTGAAGCGAAAGTAAGAGCGGCTGATGTTCAAC 1980
Db 1921 GAGATCCGGCGCTGGATAAGAACTGAAGCGAAAGTAAGAGCGGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGGCCGCTGATGCTGTCAGGGGTTATGCTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCCGCTGATGCTGTCAGGGGTTATGCTTCAAG 2040
Qy 2041 TATGAAGAACGCAAGTACGACATTAAGAAGCTGGGCGTGGATAACGCTGGCGAAAGCG 2100
Db 2041 TATGAAGAACGCAAGTACGACATTAAGAAGCTGGGCGTGGATAACGCTGGCGAAAGCG 2100
Qy 2101 GGTCTGACCTTCGCTGCTGATTAAGAAACAAACACATGATGAGACCGGATTAC 2160
Db 2101 GGTCTGACCTTCGCTGCTGATTAAGAAACAAACACATGATGAGACCGGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTTATAAGGCGAAACAGCGATGACATCAACCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTATAAGGCGAAACAGCGATGACATCAACCGCCGCTGG 2220
Qy 2221 GCATGTTCCACATCGACACGCAAGTGAATATGCTGTAACGGTATCGCGAAGCTTC 2280
Db 2221 GCATGTTCCACATCGACACGCAAGTGAATATGCTGTAACGGTATCGCGAAGCTTC 2280
Qy 2281 AAGGGTCAACATCAACACCGTTCTGTTGGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGGTCAACATCAACACCGTTCTGTTGGGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGATTCCTCGAAACCTATCTGCTGATGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGATTCCTCGAAACCTATCTGCTGATGATGAAGTCTG 2400

Qy 2401 GAACGGTTAATAAGACAAACCCCTGGTGGCTAGCGTGAAGTCTTACGAGAAAGAG 2460
Db 2401 GAACGGTTAATAAGACAAACCCCTGGTGGCTAGCGTGAAGTCTTACGAGAAAGAG 2460
Qy 2461 TTGCGGAAAGATCCACGTTATTCGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGCGGAAAGATCCACGTTATTCGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCCGCTGCGTACTCGCGTGATCAAGGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCCGCTGCGTACTCGCGTGATCAAGGCC 2580
Qy 2581 GCCAGCGTCTGTCAGATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGTCTG 2640
Db 2581 GCCAGCGTCTGTCAGATGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGTCTG 2640
Qy 2641 AACAAACAACAACAATAACAATAACAACCTCGGATCGAGGAGGATTCAGAAATTC 2700
Db 2641 AACAAACAACAACAATAACAATAACAACCTCGGATCGAGGAGGATTCAGAAATTC 2700

RESULT 13

ADO23649

ID ADO23649 standard; DNA; 7370 BP.

XX ADO23649;

XX AC ADO23649;

XX DT 01-JUL-2004 (first entry)

XX DE DNA encoding MBP-Toxop30MIX5 fusion protein.

XX KW P30 antigen; Toxo30del13C; Toxo30del12C; Toxop30 MIX1;

KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;

KW Toxoplasma gondii; ds; gene.

XX OS Toxoplasma gondii.

XX OS Synthetic.

XX PN US2004067239-A1.

XX PD 08-APR-2004.

XX PF 02-OCT-2002; 2002US-00263153.

XX PR 02-OCT-2002; 2002US-00263153.

XX PA (MAIN/) MAINE G T.

XX PA (PATE/) PATEL C B.

XX PA (GINS/) GINSBURG S R.

XX PA (BLIE/) BLIESE T R.

XX PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;

XX WI MPI; 2004-304563/28.

XX DR P-PSDB; ADO23650.

XX PT Novel purified polypeptide having sequence identity to amino acid

PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C.

PT Toxo30del12C, Toxop30 MIX1, useful for detecting presence of IgM or IgG

PT antibodies to Toxoplasma gondii.

XX Example 5; Fig 35; 114pp; English.

XX CC The invention relates to a purified P30 antigen (I) chosen from 3 fully

CC defined Toxo30del13C, Toxo30del12C and Toxop30 MIX1 sequences, having 1-6

CC amino acids added to the C-terminus of the amino acid sequence of

CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in

CC which at least one of the five C-terminal cysteine amino acids of the

CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted

CC with alanine, or comprising the amino acid sequence chosen from MBP-

CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1

CC fusion proteins. (I) is useful for detecting the presence of IgM

CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC Toxop30 fusion protein of the invention.
XX
SQ

Sequence	7370 BP;	1900 A;	1857 C;	1950 G;	1663 T;	0 U;	0 Other;
Query Match	81.5%;	Score	2690.4;	DB	12;	Length	7370;
Best Local Similarity	99.8%;	Pred. No. 0;					
Matches	2694;	Conservative	0;	Mismatches	6;	Indels	0;
Gaps	0;						
QY	1	CCGACACCATCGAATGTCGAAACCTTTCCGGTATGCGCATATGCGCCCGGAGAGA	60				
Db	1	CCGACACCATCGAATGTCGAAACCTTTCCGGTATGCGCATATGCGCCCGGAGAGA	60				
QY	61	GTCAATTGAGGTCGTAATGTAACCAAGTAACTGATATACGATGTCGACAGTATGCGG	120				
Db	61	GTCAATTGAGGTCGTAATGTAACCAAGTAACTGATATACGATGTCGACAGTATGCGG	120				
QY	121	GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCAGTTCGCGAAAA	180				
Db	121	GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCAGTTCGCGAAAA	180				
QY	181	CGCGGAAAGTGAAGCGGCGATGCGGAGCTGAATTAATTCCTCCACCGCTGGCAC	240				
Db	181	CGCGGAAAGTGAAGCGGCGATGCGGAGCTGAATTAATTCCTCCACCGCTGGCAC	240				
QY	241	AACAACTGCGGGGAAACAGTGTGCTGATTTGGGGTTGCCACCTTCAGTCTGGCCCTGC	300				
Db	241	AACAACTGCGGGGAAACAGTGTGCTGATTTGGGGTTGCCACCTTCAGTCTGGCCCTGC	300				
QY	301	ACGGCCCTGCGCAATTTGCGGGCGATTAATCTCGCGCGATCAATGCGGTGCCAGCG	360				
Db	301	ACGGCCCTGCGCAATTTGCGGGCGATTAATCTCGCGCGATCAATGCGGTGCCAGCG	360				
QY	361	TGTTGTTGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATG	420				
Db	361	TGTTGTTGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATG	420				
QY	421	TTCTCGGCAACCGCTCAGTGGGCTGATTAATTAATCTCGCTGATGATGATGATGATG	480				
Db	421	TTCTCGGCAACCGCTCAGTGGGCTGATTAATTAATCTCGCTGATGATGATGATGATG	480				
QY	481	TTGCTGTGGAAAGTCCCTGCACTAATGTTTCCGGCGTATTTCTTGATGTCCTGACCA	540				
Db	481	TTGCTGTGGAAAGTCCCTGCACTAATGTTTCCGGCGTATTTCTTGATGTCCTGACCA	540				
QY	541	CACCATCAACAGTATTAATTTCTCCATGAAGACGGTACGGACTGGCGGTGGAGCATC	600				
Db	541	CACCATCAACAGTATTAATTTCTCCATGAAGACGGTACGGACTGGCGGTGGAGCATC	600				
QY	601	TGGTCGCAATGGTTCACAGCAAACTCGCTGTTAGCGGGCCATTAAAGTCTGCTCGG	660				
Db	601	TGGTCGCAATGGTTCACAGCAAACTCGCTGTTAGCGGGCCATTAAAGTCTGCTCGG	660				
QY	661	CGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	720				
Db	661	CGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	720				
QY	721	CGGAAACGGGAGCGCATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCG	780				
Db	721	CGGAAACGGGAGCGCATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCG	780				
QY	781	ATGAGGGCATCGTTCCTCCACTGCGATGTCGATGTCGATGTCGATGTCGATGTCGATG	840				
Db	781	ATGAGGGCATCGTTCCTCCACTGCGATGTCGATGTCGATGTCGATGTCGATGTCGATG	840				
QY	841	TGCGCGCATTAACAGAGTCCGGGCTGCGGTCGATGTCGATGTCGATGTCGATGTCG	900				
Db	841	TGCGCGCATTAACAGAGTCCGGGCTGCGGTCGATGTCGATGTCGATGTCGATGTCG	900				
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAAAAGGATTTTC	960				

Db	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAAAAGGATTTTC	960				
QY	961	GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCAGCGGTGA	1020				
Db	961	GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCAGCGGTGA	1020				
QY	1021	AGGCAATCAGCTGTTGCGCGTCTCACTGCTGTAAGAAAGAAACCAACCTGCGGCCAATA	1080				
Db	1021	AGGCAATCAGCTGTTGCGCGTCTCACTGCTGTAAGAAAGAAACCAACCTGCGGCCAATA	1080				
QY	1081	CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCATTAATGACGCTGCGACACAGGTTT	1140				
Db	1081	CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCATTAATGACGCTGCGACACAGGTTT	1140				
QY	1141	CCGACCTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200				
Db	1141	CCGACCTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200				
QY	1201	GCACAAATCTCATGTTTGAAGCTTATCATGCTGACGCTGACCGTGCACCAATGCTTCTGGCG	1260				
Db	1201	GCACAAATCTCATGTTTGAAGCTTATCATGCTGACGCTGACCGTGCACCAATGCTTCTGGCG	1260				
QY	1261	TCAGGACGCAATCGAAAGCTGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1320				
Db	1261	TCAGGACGCAATCGAAAGCTGTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1320				
QY	1321	TGTCGCTCAAGCGGCACTCCCGTCTTGGATTAATGTTTTCGCGCGACATCAATACGGTT	1380				
Db	1321	TGTCGCTCAAGCGGCACTCCCGTCTTGGATTAATGTTTTCGCGCGACATCAATACGGTT	1380				
QY	1381	CTGGCAAAATCTGTAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGCTGATGCTGA	1440				
Db	1381	CTGGCAAAATCTGTAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGCTGATGCTGA	1440				
QY	1441	ATTGTGAGCGGATTAACAAATTTTTCACAGGAAACGCGCTGCTGCTGCTGCTGCTGCTG	1500				
Db	1441	ATTGTGAGCGGATTAACAAATTTTTCACAGGAAACGCGCTGCTGCTGCTGCTGCTGCTG	1500				
QY	1501	GCATTTTCAACCAAGGACCATAGATTATGAAACTGAAAGAGGTAAATCTGCTGCTGCTG	1560				
Db	1501	GCATTTTCAACCAAGGACCATAGATTATGAAACTGAAAGAGGTAAATCTGCTGCTGCTG	1560				
QY	1561	ATTAAACGCGGATTAAGGCTATAACGGTCTCGCTGAACTGCGTGAAGAAATTCGAGAAAGAT	1620				
Db	1561	ATTAAACGCGGATTAAGGCTATAACGGTCTCGCTGAACTGCGTGAAGAAATTCGAGAAAGAT	1620				
QY	1621	ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT	1680				
Db	1621	ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT	1680				
QY	1681	CGGCAACTGCGGATGCGCTGACATTAATCTTCTGGGCAACGACCGCTTTTGGTGGCTAC	1740				
Db	1681	CGGCAACTGCGGATGCGCTGACATTAATCTTCTGGGCAACGACCGCTTTTGGTGGCTAC	1740				
QY	1741	GCTCAATCTGCGCTGTTGGCTGAAATACCCCGGCAAGCGTTCCAGGACAAGCTGAT	1800				
Db	1741	GCTCAATCTGCGCTGTTGGCTGAAATACCCCGGCAAGCGTTCCAGGACAAGCTGAT	1800				
QY	1801	CGTTTACCTGGGATGCGGTACGTTTACAAACGCGAAGCTGATTTGCTTACCCGATCGCTGT	1860				
Db	1801	CGTTTACCTGGGATGCGGTACGTTTACAAACGCGAAGCTGATTTGCTTACCCGATCGCTGT	1860				
QY	1861	GAAGGTTTATCGCTGATTTATAACAAAGATCTGCTGCGGAAACCCGCCCAAAACCTGGGAA	1920				
Db	1861	GAAGGTTTATCGCTGATTTATAACAAAGATCTGCTGCGGAAACCCGCCCAAAACCTGGGAA	1920				
QY	1921	GAGATCCCGCGCTGATTAAGAACTGAAAGCAAGGTAAGAGCGCTGATGTTCAAC	1980				
Db	1921	GAGATCCCGCGCTGATTAAGAACTGAAAGCAAGGTAAGAGCGCTGATGTTCAAC	1980				
QY	1981	CTGCAGAAACCGTACTTCACTTGGCGCTGATTCGCTGACGCGGGTTATGCTTCAAG	2040				

Db 1981 CTGCAAGAACCGTACTTCACTACCTGGCGCTGATTGCTGCTGACGGGGTTATCGCTTCAAG 2040
 QY 2041 TATGAAAACGGCAAGTACGACATTAAAGACGTGGCGTGATTAAGCTGGCGCGAAAGCG 2100
 Db 2041 TATGAAAACGGCAAGTACGACATTAAAGACGTGGCGTGATTAAGCTGGCGCGAAAGCG 2100
 QY 2101 GGTCTGACCTTCTCGTGGTGAACCTGATTAAAAACAACACATGAATGACAGACACCGATTAC 2160
 Db 2101 GGTCTGACCTTCTCGTGGTGAACCTGATTAAAAACAACACATGAATGACAGACACCGATTAC 2160
 QY 2161 TCCATCGCAGAGCTGCTTTTAAATGAAGCGGAAACAGGATGACCATCAAGCGCCGTGG 2220
 Db 2161 TCCATCGCAGAGCTGCTTTTAAATGAAGCGGAAACAGGATGACCATCAAGCGCCGTGG 2220
 QY 2221 GCATGGTCCAAATCGACACACAGCAAGTGAATTATGGTGAACCGTACTGCGGACCTTC 2280
 Db 2221 GCATGGTCCAAATCGACACACAGCAAGTGAATTATGGTGAACCGTACTGCGGACCTTC 2280
 QY 2281 AAGGTCACCATCCAAACCGTTCTGGCGGTGCTGAGCGAGGTATTAAACGCCGCCAGT 2340
 Db 2281 AAGGTCACCATCCAAACCGTTCTGGCGGTGCTGAGCGAGGTATTAAACGCCGCCAGT 2340
 QY 2341 CGAACAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGGTTCTG 2400
 Db 2341 CGAACAAGAGCTGGCAAAAGAGTTCTCGAAACTATCTGCTGACTGATGAAGGTTCTG 2400
 QY 2401 GAAGCGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAGAG 2460
 Db 2401 GAAGCGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAGAG 2460
 QY 2461 TTGGCGAAGATCCACGTTATTGCGGCCACCATGGAAGACGCCCAAGAGGTGAATCATG 2520
 Db 2461 TTGGCGAAGATCCACGTTATTGCGGCCACCATGGAAGACGCCCAAGAGGTGAATCATG 2520
 QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTGATGTCGCTGCTGCTGCTGCTGCTGCTGCTG 2580
 Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTGATGTCGCTGCTGCTGCTGCTGCTGCTGCTG 2580
 QY 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGACGCCGACAGACTAATTCGAGCTCG 2640
 Db 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGACGCCGACAGACTAATTCGAGCTCG 2640
 QY 2641 AACAAACAAACAAATACAAATAACAAACAACTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
 Db 2641 AACAAACAAACAAATACAAATAACAAACAACTCGGGATCGAGGGAAGGATTCAGAAATTC 2700

RESULT 14
 ADO23644
 ID ADO23644 standard; DNA; 7370 BP.
 AC ADO23644;
 XX ADO23644;
 DT 01-JUL-2004 (first entry)
 XX DNA encoding MBP-Toxop30MIX3 fusion protein.
 DE
 DE
 XX P30 antigen; Toxo30del13C; Toxo30del12C; Toxo30 MIX1;
 KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
 KW Toxoplasma gondii; ds; gene.
 XX
 OS Toxoplasma gondii.
 OS Synthetic.
 XX
 XX US2004067239-A1.
 XX
 XX 08-APR-2004.
 XX
 XX 02-OCT-2002; 2002US-00263153.
 XX
 XX 02-OCT-2002; 2002US-00263153.
 XX
 XX (MAIN/) MAINS G T.

PA (PATE/) PATEL C B.
 PA (GINS/) GINSBURG S R.
 PA (BLIE/) BLIESE T R.
 XX
 FI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
 XX
 XX WPI; 2004-304563/28.
 DR P-PSDB; ADO23645.
 XX
 PT Novel purified polypeptide having sequence identity to amino acid
 PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
 PT Toxo30del12C, Toxo30 MIX1, useful for detecting presence of IgM or IgG
 PT antibodies to Toxoplasma gondii.
 XX
 XX Example 5; Fig 33; 114pp; English.

The invention relates to a purified P30 antigen (I) chosen from 3 fully defined Toxo30del13C, Toxo30del12C and Toxo30 MIX1 sequences, having 1-6 amino acids added to the C-terminus of the amino acid sequence of Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in which at least one of the five C-terminal cysteine amino acids of the amino acid sequence of Toxo30del13C P30 antigen sequence is substituted with alanine, or comprising the amino acid sequence chosen from MBP-Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1 fusion proteins. (I) is useful for detecting the presence of IgM antibodies to Toxoplasma gondii in a test sample, which involves contacting the test sample suspected of containing the IgM antibodies with a composition comprising (I) and detecting the presence of (I)/IgM antibody complexes. The present sequence represents DNA encoding a MBP-Toxo30 fusion protein of the invention.

Sequence 7370 BP; 1900 A; 1857 C; 1950 G; 1663 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGCATGATAGCCGCCGGAAGAGA 60
 Db 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGCATGATAGCCGCCGGAAGAGA 60
 QY 61 GTCATTCAGGGTGGTGAATGTGAACCAAGTAACTATACATGTCGAGAGTATGCGG 120
 Db 61 GTCATTCAGGGTGGTGAATGTGAACCAAGTAACTATACATGTCGAGAGTATGCGG 120
 QY 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACCAAGTAACTATGTCGAGAGTATGCGG 180
 Db 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACCAAGTAACTATGTCGAGAGTATGCGG 180
 QY 181 CGCGGAAAAAGTGAAGCGGCGATGGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
 Db 181 CGCGGAAAAAGTGAAGCGGCGATGGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
 QY 241 AACACTGGCGGCAAAAGTGTGCTGATGAGGCTTGCACCTTCCAGTCTGGCCCTGC 300
 Db 241 AACACTGGCGGCAAAAGTGTGCTGATGAGGCTTGCACCTTCCAGTCTGGCCCTGC 300
 QY 301 AGCGCGCTCGCAAAATTTGTCGGGCGATTAATATCTCGGCGGATCACTGAGTGGTGGCAGCG 360
 Db 301 AGCGCGCTCGCAAAATTTGTCGGGCGATTAATATCTCGGCGGATCACTGAGTGGTGGCAGCG 360
 QY 361 TGGTGGTGTGATGGTGAACCAAGCGGCTGAAAGCGCTGTAAAGCGGCGGTGCACATC 420
 Db 361 TGGTGGTGTGATGGTGAACCAAGCGGCTGAAAGCGCTGTAAAGCGGCGGTGCACATC 420
 QY 421 TTCTGCGCAACCGCTCAGTGGCTGATCATTAATATCCGCTGGATGACAGGATGCCA 480
 Db 421 TTCTGCGCAACCGCTCAGTGGCTGATCATTAATATCCGCTGGATGACAGGATGCCA 480
 QY 481 TTGCTGTGGAAGCTGCTGCACTTAATTTTCCGGCGTTATTTCTTGTATGTCCTGACCGAGA 540
 Db 481 TTGCTGTGGAAGCTGCTGCACTTAATTTTCCGGCGTTATTTCTTGTATGTCCTGACCGAGA 540

QY 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACCGGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACCGGACTGGCGGTGGAGCATC 600
QY 601 TGGTCGCATTGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Db 601 TGGTCGCATTGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGCTGTGGCTGGCGTGAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTGTGGCTGGCGTGAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGACATGGAGTGCATGTCCGGTTTCAACAAACCATGCAATGCTGA 780
Db 721 CGGAACGGGAAGGACATGGAGTGCATGTCCGGTTTCAACAAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAAGATCAGATGGCGCTGGCGCGAA 840
Db 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTTGCCAAGATCAGATGGCGCTGGCGCGAA 840
QY 841 TGGCGCCCATACCAAGTCCGGGCTGGCGGTTGGTGGCGATATCTCGGTAGTGGGATAG 900
Db 841 TGGCGCCCATACCAAGTCCGGGCTGGCGGTTGGTGGCGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACAGCTGGACCGCTTCTGCAACTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACAGCTGGACCGCTTCTGCAACTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGGCCGCTCACTGGTGAAGAAAGAAACCACTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCGCTCACTGGTGAAGAAAGAAACCACTGGCGGCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCCGATTCATTAATGCACTGCGACGACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCCGATTCATTAATGCACTGCGACGACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAATTTCTCATGTTTGAACGCTTATCATCGACTGCACGCTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTTCTCATGTTTGAACGCTTATCATCGACTGCACGCTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGCAGGCATCGGAAGCTGTGGTATGGCTGTGCAAGTTCGTAATCACTGCATTAATTCG 1320
Db 1261 TCAGGCAGGCATCGGAAGCTGTGGTATGGCTGTGCAAGTTCGTAATCACTGCATTAATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGTCTGGATATGTTTTTGGCGCGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTGGATATGTTTTTGGCGCGACATCATACGGTT 1380
QY 1381 CTGGCAAAATATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTGGA 1440
Db 1381 CTGGCAAAATATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTGGA 1440
QY 1441 ATTGTAGCGGATACAAATTTTCAACAGAAACAGCCAGTCCGTTAGGTGTTTTCAAGA 1500
Db 1441 ATTGTAGCGGATACAAATTTTCAACAGAAACAGCCAGTCCGTTAGGTGTTTTCAAGA 1500
QY 1501 GCACCTTCCACCAAGGACCATAGATTGAAAATCTGAAGAGGTAAATCTGGTATCTGG 1560
Db 1501 GCACCTTCCACCAAGGACCATAGATTGAAAATCTGAAGAGGTAAATCTGGTATCTGG 1560
QY 1561 ATTTAAACGGGATAAAGGCTTAAACGGTCTCGCTGAAAGTCGGTAAAGAAATTCAGAAAGAT 1620
Db 1561 ATTTAAACGGGATAAAGGCTTAAACGGTCTCGCTGAAAGTCGGTAAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAAATTAAGTCAACGGTTGAGCATCCGGATAAACTGGAAAGAGAAATTTCCACAGGTT 1680

Db 1621 ACCGGAAATTAAGTCAACGGTTGAGCATCCGGATAAACTGGAAAGAGAAATTTCCACAGGTT 1680
QY 1681 CGCGCAACTGCGGATGGCCCTTGACATATCTTCTTGGGCACACGACCGCTTTTGGTGGCTAC 1740
Db 1681 CGCGCAACTGCGGATGGCCCTTGACATATCTTCTTGGGCACACGACCGCTTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCGGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTGGCTGAAATCAACCCGGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGGCGTACGTTTACAAACGGCAAGCTGATTTGCTTACCCGATCCCTGT 1860
Db 1801 CCGTTTACCTGGGATGGCGTACGTTTACAAACGGCAAGCTGATTTGCTTACCCGATCCCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAGCGAAAGSTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAGCGAAAGSTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTTCACCTGGCGCTGATTTGCTGTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGGCGCTGATTTGCTGTGACGGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAACAAACATGAAATGCAAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAACAAACATGAAATGCAAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCAACATCGACACACGAGCAAAAGTGAATTAATGCTGTAACCGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACACGAGCAAAAGTGAATTAATGCTGTAACCGTACTGCGGACCTTC 2280
QY 2281 AAGGTCACCAATCCAAACCGTTGCTGGCGTGTGAGCGCAGGATTAATAACCGCCAGT 2340
Db 2281 AAGGTCACCAATCCAAACCGTTGCTGGCGTGTGAGCGCAGGATTAATAACCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
Db 2401 GAAGCGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCACGATTTTGGCGCCACCATGGAAGAAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTTTGGCGCCACCATGGAAGAAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGGACACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGGACACTAAATTCGAGCTCG 2640
QY 2641 AACCAACAAACAAATAACAAATAACAACTTCGGGATCGAGGGAAGGATTTCAAAATTC 2700
Db 2641 AACCAACAAACAAATAACAAATAACAACTTCGGGATCGAGGGAAGGATTTCAAAATTC 2700

AD023594
ID AD023594 standard; DNA; 7403 BP.
XX
AC AD023594;
XX
DT 01-JUL-2004 (first entry)
XX
DE DNA encoding MBP-Toxop30del12(52-311aa) fusion protein.
XX
KW P30 antigen; Toxop30del13C; Toxop30del12C; Toxop30 MIX1;
KW MBP-Toxop30del13C(52-300aa); MBP-Toxop30del14C(52-294aa); MBP-Toxop30MIX1;
KW Toxoplasma gondii; da; gene.
XX
OS Toxoplasma gondii.
OS Synthetic.
XX
FN US2004067239-A1.
XX
PD 08-APR-2004.
XX
PF 02-OCT-2002; 2002US-00263153.
XX
PR 02-OCT-2002; 2002US-00263153.
XX
PA (MAIN/) MAINE G T.
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX
PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX
DR WPI; 2004-304563/28.
DR P-PSDB; AD023595.
XX
PT Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxop30del13C,
PT Toxop30del12C, Toxop30 MIX1, useful for detecting presence of IgM or IgG
PT antibodies to Toxoplasma gondii.
XX
PS Example 2; Fig 11; 114pp; English.
XX
CC The invention relates to a purified P30 antigen (I) chosen from 3 fully
CC defined Toxop30del13C, Toxop30del12C and Toxop30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxop30del12C P30 antigen sequence, or comprising an amino acid sequence in
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxop30del13C P30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxop30del13C(52-300aa), MBP-Toxop30del14C(52-294aa) and MBP-Toxop30MIX1
CC fusion proteins. (I) is useful for detecting the presence of IgM
CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC Toxop30 fusion protein of the invention.
SQ Sequence 7403 BP; 1907 A; 1863 C; 1961 G; 1672 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7403;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGTTGAAATGTGAACACCTTTCGCGGTATGGCATATGCGCCCGGAAGAGA 60
DB 1 CCGACACCATCGAATGTTGAAATGTGAACACCTTTCGCGGTATGGCATATGCGCCCGGAAGAGA 60
QY 61 GTCAATTACAGGTGGTGAATGTGAACACCTTTCGCGGTATGGCATATGCGCCCGGAAGAGA 120
DB 61 GTCAATTACAGGTGGTGAATGTGAACACCTTTCGCGGTATGGCATATGCGCCCGGAAGAGA 120
QY 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACACCGCCAGCCAGCGTTTCTCGGAAA 180
DB 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACACCGCCAGCCAGCGTTTCTCGGAAA 180

QY 181 CGCGGAAAAAGTGAAGCGGATGGCGGAGCTGAATTAATTCCTCCAAACCGGTGGCAC 240
DB 181 CGCGGAAAAAGTGAAGCGGAGCTGGCGGAGCTGAATTAATTCCTCCAAACCGGTGGCAC 240
QY 241 AACAACTGGCGGCAAAACAGTGTGTGTGATGGCGGTGGCCACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGCAAAACAGTGTGTGTGATGGCGGTGGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTCCGCAAAATTTGTCCGCGGATTAATCTCCGCGCGATCAATCGGGTCCAGCG 360
DB 301 ACGCGCGTCCGCAAAATTTGTCCGCGGATTAATCTCCGCGCGATCAATCGGGTCCAGCG 360
QY 361 TGGTGGTGTGATGTAGAACGAAGCGCGTCCGAAGCTGTGAAGCGCGGTGCACAAATC 420
DB 361 TGGTGGTGTGATGTAGAACGAAGCGCGTCCGAAGCTGTGAAGCGCGGTGCACAAATC 420
QY 421 TTCTCGCGCAACCGGTCAAGTGGCTGATCAATTAATCTATCCGCTGGATGACCAAGTCCCA 480
DB 421 TTCTCGCGCAACCGGTCAAGTGGCTGATCAATTAATCTATCCGCTGGATGACCAAGTCCCA 480
QY 481 TTGCTGTGGAAGTCCCTGCACATAATTTCCGCGGTATTTCTTGATGTCTTCGACGAGA 540
DB 481 TTGCTGTGGAAGTCCCTGCACATAATTTCCGCGGTATTTCTTGATGTCTTCGACGAGA 540
QY 541 CACCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
DB 541 CACCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
QY 601 TGGTCGCAATTTGGGTCAACAGCAATCGCTGTAGCGGCGCATTAAGTCTGTCTCGG 660
DB 601 TGGTCGCAATTTGGGTCAACAGCAATCGCTGTAGCGGCGCATTAAGTCTGTCTCGG 660
QY 661 CGGCTGTGCTGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGGCTGTGCTGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGCGGACTGGAGTGCATGTCGGTTCGATTTTCAACAAACCATGCAAAATGCTGA 780
DB 721 CGGAACGGGAAGCGGACTGGAGTGCATGTCGGTTCGATTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGCAATCGTTCCCACTGGGATGCTGGTGGCCAAAGATCAGATGGCGTGGCGCAA 840
DB 781 ATGAGGCAATCGTTCCCACTGGGATGCTGGTGGCCAAAGATCAGATGGCGTGGCGCAA 840
QY 841 TCGCGCCATTAACGAGTCCGCGTGGCGTGGTGGCGGATATCTCGGTAGTGGGATAG 900
DB 841 TCGCGCCATTAACGAGTCCGCGTGGCGTGGTGGCGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAACAGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAACAGATTTTC 960
QY 961 GCCTGCTGGGCAAAACAGCGGTGGACCGTTCGTCGCAACTCTCTCAGGCGCGCGGTGA 1020
DB 961 GCCTGCTGGGCAAAACAGCGGTGGACCGTTCGTCGCAACTCTCTCAGGCGCGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTGCGCGTCTCACTGGTGAAGAAAGAAACCCCTGGCGCCCAATA 1080
DB 1021 AGGCAATCAGCTGTGCGCGTCTCACTGGTGAAGAAAGAAACCCCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTGGCCGATTCATTAATGACAGTGGCAGCAGAGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGGTGGCCGATTCATTAATGACAGTGGCAGCAGAGTTT 1140
QY 1141 CCCGACTGGAAAGCGGCGAGTGGCAAGCAATTAATGTAGTTAGTCACTCAATTAG 1200
DB 1141 CCCGACTGGAAAGCGGCGAGTGGCAAGCAATTAATGTAGTTAGTCACTCAATTAG 1200
QY 1201 GCACAAATCTCATGTTTGAAGCTTATCATCGACTGCGCGGTGCACCAATGCTTCTGGCG 1260
DB 1201 GCACAAATCTCATGTTTGAAGCTTATCATCGACTGCGCGGTGCACCAATGCTTCTGGCG 1260

QY 1261 TCAGGAGCCATCGAAGCTGTGGTATGGCTGTGCAGGTGCTTAATCACTGCATAATTCTG 1320
Db 1261 TCAGGAGCCATCGAAGCTGTGGTATGGCTGTGCAGGTGCTTAATCACTGCATAATTCTG 1320
QY 1321 TGTGCTCAAGCGCACCTCCCGTTCTCGATATATGTTTTTGGCCGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACCTCCCGTTCTCGATATATGTTTTTGGCCGACATCATACGGTT 1380
QY 1391 CTGGCAATATTTGAAATGAGCTTGTGACATTAATCATCGGCTCGTATATGTTGTGA 1440
Db 1391 CTGGCAATATTTGAAATGAGCTTGTGACATTAATCATCGGCTCGTATATGTTGTGA 1440
QY 1441 ATTGTGAGCGATACAAATTTACACAGGAACAGCCAGTCCGTTAGGTTGTTTCAGA 1500
Db 1441 ATTGTGAGCGATACAAATTTACACAGGAACAGCCAGTCCGTTAGGTTGTTTCAGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAAACTGAAAGAGTAAACTGGTATCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGCATATGAAATTCGAAGAGTAAACTGGTATCTGG 1560
QY 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACGGTTGAGCATCGGATTAACCTGGAAGAGAAATTTCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGGTTGAGCATCGGATTAACCTGGAAGAGAAATTTCCACAGGTT 1680
QY 1681 GCGGCAATGCGGATGCGCTGACATTTATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAATGCGGATGCGCTGACATTTATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTGGCTGAATATCAACCGCAAGCGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTGGCTGAATATCAACCGCAAGCGTTCCAGGACAAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCGGTAGCTTACACCGCAAGCTGATTGCTTACCGGATCGCTGT 1860
Db 1801 CCGTTTACCTGGATGCGGTAGCTTACACCGCAAGCTGATTGCTTACCGGATCGCTGT 1860
QY 1861 GAAGGTTTATCGCTGATTTATAAAGATCTGCTGCCGAACCCGCCAAACCTGGGAA 1920
Db 1861 GAAGGTTTATCGCTGATTTATAAAGATCTGCTGCCGAACCCGCCAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGGGTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGGGTTATCGGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGGCGTGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGGCGTGATTAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGTTGACCTGATTAAAAACAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTGACCTGATTAAAAACAACACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCAACATCGACACGCAAGTGAATATGTTGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACGCAAGTGAATATGTTGTAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCAAAACCGTTTGGTGGCTGCTGAGCGCAGGTATTAAACGCCCGCAGT 2340
Db 2281 AAGGGTCAACCATCAAAACCGTTTGGTGGCTGCTGAGCGCAGGTATTAAACGCCCGCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400

Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAACGGTTAATAAAGACAAACCCCTGGGTCCGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAACGGTTAATAAAGACAAACCCCTGGGTCCGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAGATCCACGTATTTGCCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGTATTTGCCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGA CGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGAGCGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAAATAACATAACAAACCTCGGATCGAGGGAAGGATTTCAGAATTC 2700
Db 2641 AACAAACAAACAAATAACATAACAAACCTCGGATCGAGGGAAGGATTTCAGAATTC 2700

Search completed: August 1, 2005, 04:28:59
Job time : 1101.6 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 352.4 Seconds
(without alignments)
15322.682 Million cell updates/sec

Title: US-09-765-555B-18

Perfect score: 3300

Sequence: 1 ccgacacatcgatgtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2700	81.8	7475	2	US-08-971-036-1
2	2700	81.8	7475	3	US-09-096-570-1
3	2700	81.8	7475	3	US-09-265-617B-1
c	1417.6	43.0	5926	3	US-09-027-169-3
5	1362	41.3	5201	4	US-09-640-882-2
6	1362	41.3	5201	4	US-09-640-882-3
7	1204	36.5	4557	4	US-08-778-717-5
8	1200.8	36.4	3832	1	US-08-148-675A-2
c	1199.2	36.3	5248	3	US-08-487-283A-18
c	1199.2	36.3	5248	5	PCT-US96-05611A-21
c	1119.2	36.3	5312	4	US-10-263-103-35
c	1119.2	36.3	5443	2	US-08-929-967-1
c	1119.2	36.3	5502	4	US-09-702-705-785
c	1119.2	36.3	5502	4	US-09-736-457-785
c	1119.2	36.3	5502	4	US-09-614-124B-785
c	1119.2	36.3	5502	4	US-09-671-325-785
c	1119.2	36.3	5502	4	US-09-589-184-785
c	1119.2	36.3	5502	4	US-09-658-824-785
c	1119.2	36.3	5502	4	US-08-929-967-3
c	1119.2	36.3	5873	4	US-09-695-437A-62
c	1119.2	36.3	6353	4	US-09-702-705-784
c	1119.2	36.3	6353	4	US-09-736-457-784
c	1119.2	36.3	6353	4	US-09-614-124B-784
c	1119.2	36.3	6353	4	US-09-671-325-784
c	1119.2	36.3	6353	4	US-09-589-184-784
c	1119.2	36.3	6353	4	US-09-658-824-784
c	1119.2	36.3	6361	3	US-08-646-538-7

28 1199.2 36.3 6361 3 US-09-503-222-7
c 29 1199.2 36.3 6727 2 US-08-125-462-2
c 30 1199.2 36.3 6727 2 US-08-891-848-2
c 31 1199.2 36.3 6799 2 US-08-125-462-5
c 32 1199.2 36.3 6799 2 US-08-891-848-5
c 33 1199.2 36.3 7659 3 US-09-128-314-4
c 34 1199.2 36.3 7676 3 US-09-056-556-213
c 35 1199.2 36.3 7676 3 US-09-072-596-208
c 36 1199.2 36.3 7676 4 US-09-072-967-213
c 37 1199.2 36.3 7676 4 US-09-287-849-9
c 38 1199.2 36.3 8031 3 US-09-643-597-254
c 39 1199.2 36.3 8031 4 US-09-480-884A-254
c 40 1199.2 36.3 8031 4 US-09-542-615A-254
c 41 1199.2 36.3 8031 4 US-09-606-421B-254
c 42 1199.2 36.3 8031 4 US-09-476-496A-254
c 43 1199.2 36.3 8031 4 US-09-630-940B-254
c 44 1199.2 36.3 8157 3 US-09-128-314-3
c 45 1199.2 36.3 8501 3 US-08-793-900-1

ALIGNMENTS

RESULT 1

US-08-971-036-1
; Sequence 1, Application US/08971036
; Patent No. 5866684
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R
; APPLICANT: Hurst, David N
; APPLICANT: Jones, Philip S
; APPLICANT: Kay, Paul B
; APPLICANT: Raynham, Tony M
; APPLICANT: Wilson, Francis X
; TITLE OF INVENTION: Amino Acid Derivatives
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: N.J.
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/971.036
; FILING DATE: 14-NOV-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER: GB 9623908.2
; FILING DATE: 18-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kreisler, Lewis J
; REGISTRATION NUMBER: 38522
; REFERENCE/DOCKET NUMBER: RAN 4430/073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (973) 235-4387
; TELEFAX: (973) 235-2363
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7475 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
US-08-971-036-1

Query Match 81.8%; Score 2700; DB 2; Length 7475;

Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 CGACACCATCAATGCTCAAAACCTTTCGGCGGTATGCGCATGATAGCGCCCGGAAGAGA 60
Db 1 CGACACCATCAATGCTCAAAACCTTTCGGCGGTATGCGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTCAGGGTGGTGAATGTGAACACCAAGTAACGTTATACGATGTCGACAGTATGCCG 120
Db 61 GTCAATTCAGGGTGGTGAATGTGAACACCAAGTAACGTTATACGATGTCGACAGTATGCCG 120
QY 121 GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACAGGCCAGCCAGCGTTTCTCGGAAAA 180
Db 121 GTGTCCTTATCAGACCGTTTCCCGCGTGGTGAACAGGCCAGCCAGCGTTTCTCGGAAAA 180
QY 181 CCGCGGAAAAAGTGAAGCGGGATGGCGGAGCTGAATTAACATTCACACCGCTGGGCAC 240
Db 181 CCGCGGAAAAAGTGAAGCGGGATGGCGGAGCTGAATTAACATTCACACCGCTGGGCAC 240
QY 241 AACAACTGGCGGCAAAACAGTGGTGTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAAAACAGTGGTGTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTCCGCAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCGACG 360
Db 301 ACGCGCGTCCGCAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCGACG 360
QY 361 TGGTGGTTCGATGGTGAACGAAGCGGCTGGAAGCGCTGTAAGCGCGGTGACCAATC 420
Db 361 TGGTGGTTCGATGGTGAACGAAGCGGCTGGAAGCGCTGTAAGCGCGGTGACCAATC 420
QY 421 TTCTCGCGCAACGCGTCAAGTGGGTGATCAATACTATCCGTCGATGACCAAGATGCCA 480
Db 421 TTCTCGCGCAACGCGTCAAGTGGGTGATCAATACTATCCGTCGATGACCAAGATGCCA 480
QY 481 TTGCTGTGGAAGCTCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTGACGAGA 540
Db 481 TTGCTGTGGAAGCTCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCCTGACGAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGCACTGGGCGTGGAGATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGCACTGGGCGTGGAGATC 600
QY 601 TGGTGCAATGGGTCAACAGCAATTCGCGCTGTTAGCGGGCCCAATTAAGTTCCTCTCGG 660
Db 601 TGGTGCAATGGGTCAACAGCAATTCGCGCTGTTAGCGGGCCCAATTAAGTTCCTCTCGG 660
QY 661 CCGCTCTGCGTCTGCGTGGCATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CCGCTCTGCGTCTGCGTGGCATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCGACTGGAGTGCCATGTCGGTTTTTCAACAAACCATGCAATGCTGA 780
Db 721 CGGAACGGGAAGGCGACTGGAGTGCCATGTCGGTTTTTCAACAAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTCGCATGCTGTTGTCGAACGATCAGATGGCGTGGCGCAA 840
Db 781 ATGAGGGCATGTTCCCACTCGCATGCTGTTGTCGAACGATCAGATGGCGTGGCGCAA 840
QY 841 TGGCGGCATTAACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGGCATTAACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTTAAACCAATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTTAAACCAATCAAAACAGGATTTTC 960
QY 961 GCCTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGTGA 1020
Db 961 GCCTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGTGA 1020
QY 1021 AGGCAATACGCTGTTCCCGTCTCACTGGTGAAGAAAAACACCCCTGGCGCCCAATA 1080
Db 1021 AGGCAATACGCTGTTCCCGTCTCACTGGTGAAGAAAAACACCATGAATGCAGACCCGATTAC 2160
```

```
Db 1021 AGGCAATACGCTGTTCCCGTCTCACTGGTGAAGAAAAACACCCCTGGCGCCCAATA 1080
QY 1081 CGAAAAACCGCTCTCCCGCGCGTGGCGGATTCATTAATGCACTGGCACGACGAGTTT 1140
Db 1081 CGAAAAACCGCTCTCCCGCGCGTGGCGGATTCATTAATGCACTGGCACGACGAGTTT 1140
QY 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTAG 1200
Db 1141 CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTAG 1200
QY 1201 GCACAAATCTCATGTTTTGACAGCTTATCATCGACTGCACGCTGCACCAATGCTTCTGCGG 1260
Db 1201 GCACAAATCTCATGTTTTGACAGCTTATCATCGACTGCACGCTGCACCAATGCTTCTGCGG 1260
QY 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGTCAGGTCGTAAATCACTGCATAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGTCAGGTCGTAAATCACTGCATAATTCG 1320
QY 1321 TGTGCTCAAGGGCGCACTCCCGTTCTGATAAATGTTTTTTCGCGCGACATCAAAACGGTT 1380
Db 1321 TGTGCTCAAGGGCGCACTCCCGTTCTGATAAATGTTTTTTCGCGCGACATCAAAACGGTT 1380
QY 1381 CTGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGGA 1440
Db 1381 CTGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGGA 1440
QY 1441 ATTGTGACCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTATAGTGTGTTTACGA 1500
Db 1441 ATTGTGACCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTATAGTGTGTTTACGA 1500
QY 1501 GCACCTTCAACAAACAGGACCATAGATTTATGAAACTGGAAGAGTAACTGTTGTAATCTGG 1560
Db 1501 GCACCTTCAACAAACAGGACCATAGATTTATGAAACTGGAAGAGTAACTGTTGTAATCTGG 1560
QY 1561 ATTAACCGCGGATTAAGGCTATAACGGTCTCGTGAAGTGGTAAAGAAATTCGAGAAGAT 1620
Db 1561 ATTAACCGCGGATTAAGGCTATAACGGTCTCGTGAAGTGGTAAAGAAATTCGAGAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAGAGAAATTCACACAGTTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAGAGAAATTCACACAGTTT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTTCTCTGGGCACACGACCGCTTTGGTGCGTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTTCTCTGGGCACACGACCGCTTTGGTGCGTAC 1740
QY 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCAACCCCGGAACAAAGCGTTCCAGGACAACTGTAT 1800
Db 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCAACCCCGGAACAAAGCGTTCCAGGACAACTGTAT 1800
QY 1801 CCGTTTTACCTGGGATCCCGTACGTTACAAACGGCAAGCTGATTCGTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTTACCTGGGATCCCGTACGTTACAAACGGCAAGCTGATTCGTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAAACCCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAAACCCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGATAAGAACTGAAGAGGTAAGGTAAGAGCGCGCTGATGTTTCAAC 1980
Db 1921 GAGATCCCGCGCTGATAAGAACTGAAGAGGTAAGGTAAGAGCGCGCTGATGTTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTCGTTGACGGGGGTTATGCGTTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTCGTTGACGGGGGTTATGCGTTTCAAG 2040
QY 2041 TATGAAGAACCGCAAGTACGACATTAAGACGCTGGCGTGGATAAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAGAACCGCAAGTACGACATTAAGACGCTGGCGTGGATAAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAAGAAACCAACATCAATGCAGACCCGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAAGAAACCAACATCAATGCAGACCCGATTAC 2160
```


QY 841 TGCGGCCATTACCGAGTCCGGCTCGCGTTGGTGCAGATATCTCGTGTAGTGGGATACG 900
DB |||||
QY 841 TGCGGCCATTACCGAGTCCGGCTCGCGTTGGTGCAGATATCTCGTGTAGTGGGATACG 900
DB |||||
QY 901 ACGATACCGAAGACAGCTCATGTATTATCCGCCCTTAAACCAATCAAAACAGGATTTTC 960
DB |||||
QY 901 ACGATACCGAAGACAGCTCATGTATTATCCGCCCTTAAACCAATCAAAACAGGATTTTC 960
DB |||||
QY 961 GCCTCTCGGGCAACACGAGCTGGAGCCGCTTCTGTCGAATCTCTCAGAGCCAGCGGTGA 1020
DB |||||
QY 961 GCCTCTCGGGCAACACGAGCTGGAGCCGCTTCTGTCGAATCTCTCAGAGCCAGCGGTGA 1020
DB |||||
QY 1021 AGGCAATCAGCTGTGTGCCCTCTCACTGGTGAAGAAACACACCTTGGCGCCCAATA 1080
DB |||||
QY 1021 AGGCAATCAGCTGTGTGCCCTCTCACTGGTGAAGAAACACACCTTGGCGCCCAATA 1080
DB |||||
QY 1081 CGCAACCGCTCTCCCGCGGTGGCGGATTCATTAATCGAGTGGCAGCAGGTTT 1140
DB |||||
QY 1081 CGCAACCGCTCTCCCGCGGTGGCGGATTCATTAATCGAGTGGCAGCAGGTTT 1140
DB |||||
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTCACCTCATTTAG 1200
DB |||||
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTCACCTCATTTAG 1200
DB |||||
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCACGCTGACCAATCTTCTGGCG 1260
DB |||||
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCACGCTGACCAATCTTCTGGCG 1260
DB |||||
QY 1261 TCAGCAGCCATCGGAGCTGTGATGCTGTGCGAGTGTGATCACTGCTGATTAATTCG 1320
DB |||||
QY 1261 TCAGCAGCCATCGGAGCTGTGATGCTGTGCGAGTGTGATCACTGCTGATTAATTCG 1320
DB |||||
QY 1321 TGTGCTCAAGCGGCACTCCCGTCTGTGATAATGTTTTGCGCGGACATCAACCGTT 1380
DB |||||
QY 1321 TGTGCTCAAGCGGCACTCCCGTCTGTGATAATGTTTTGCGCGGACATCAACCGTT 1380
DB |||||
QY 1381 CTGGCAATATCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
DB |||||
QY 1381 CTGGCAATATCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATAATGTGGA 1440
DB |||||
QY 1441 ATTGTAGCGGATACAAATTTCAACAGGAACAGCCAGTCCGTTAGTGTGTTTACGA 1500
DB |||||
QY 1441 ATTGTAGCGGATACAAATTTCAACAGGAACAGCCAGTCCGTTAGTGTGTTTACGA 1500
DB |||||
QY 1501 GCATTTCAACCAAGGACCAATAGATTATGAAACTGAAGAGTAACTGGTAATCTGG 1560
DB |||||
QY 1501 GCATTTCAACCAAGGACCAATAGATTATGAAACTGAAGAGTAACTGGTAATCTGG 1560
DB |||||
QY 1561 ATTTAACCGGATTAAGGCTATAACGGTCTCGCTGAAGTTCGGTAAGAAATTCGAGAAAGAT 1620
DB |||||
QY 1561 ATTTAACCGGATTAAGGCTATAACGGTCTCGCTGAAGTTCGGTAAGAAATTCGAGAAAGAT 1620
DB |||||
QY 1621 ACCGGAATTAAGTCAACGTTGAGCTCCGATTAACCTGGAAGAGAAATTCGCCAGGTT 1680
DB |||||
QY 1621 ACCGGAATTAAGTCAACGTTGAGCTCCGATTAACCTGGAAGAGAAATTCGCCAGGTT 1680
DB |||||
QY 1681 GCGCAACTGCGGATGCGCTGACATATTCTTTGGGCACACGACCCGTTTGGTGGCTAC 1740
DB |||||
QY 1681 GCGCAACTGCGGATGCGCTGACATATTCTTTGGGCACACGACCCGTTTGGTGGCTAC 1740
DB |||||
QY 1741 GCTCAATCTGGCTGTGGCTGAATCAACCCGACAAAGGTTTCCAGGCAAGCTGTAT 1800
DB |||||
QY 1741 GCTCAATCTGGCTGTGGCTGAATCAACCCGACAAAGGTTTCCAGGCAAGCTGTAT 1800
DB |||||
QY 1801 CCGTTTACCTGGGATGCGCTGAGTTCAACCGCAAGCTGATTTGCTTACCCGATGCTGTT 1860
DB |||||
QY 1801 CCGTTTACCTGGGATGCGCTGAGTTCAACCGCAAGCTGATTTGCTTACCCGATGCTGTT 1860
DB |||||
QY 1861 GAAGCGTTATCGTATTATTAACAAAGATCTGCTGCCGAAACCCGCAAAACCTGGGAA 1920
DB |||||
QY 1861 GAAGCGTTATCGTATTATTAACAAAGATCTGCTGCCGAAACCCGCAAAACCTGGGAA 1920
DB |||||

QY 1921 GAGATCCGGCGCTGATATAAGAACTGAAAGCGAAAGTGAAGCGCGCTGATCTTCAAC 1980
DB |||||
QY 1921 GAGATCCGGCGCTGATATAAGAACTGAAAGCGAAAGTGAAGCGCGCTGATCTTCAAC 1980
DB |||||
QY 1981 CTGCAAGAACCGTACTTTCACCTGCGCTGATTTGCTGCTGACGGGGTATTCGCTTCAAG 2040
DB |||||
QY 1981 CTGCAAGAACCGTACTTTCACCTGCGCTGATTTGCTGCTGACGGGGTATTCGCTTCAAG 2040
DB |||||
QY 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGCTGGATTAACGCTGGCGGAAAGCG 2100
DB |||||
QY 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGCTGGATTAACGCTGGCGGAAAGCG 2100
DB |||||
QY 2101 GGTCTGACCTTCTGCTGCTGCTGATTTAAACAAACACATGATGATGAGACCGGATTC 2160
DB |||||
QY 2101 GGTCTGACCTTCTGCTGCTGCTGATTTAAACAAACACATGATGATGAGACCGGATTC 2160
DB |||||
QY 2161 TCCATCCGAAAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCGCCGTGG 2220
DB |||||
QY 2161 TCCATCCGAAAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCGCCGTGG 2220
DB |||||
QY 2221 GCATGCTCCAACTCGACACAGCAAGTGAATTAATGCTGTAACGCTGCTGCGACCTTC 2280
DB |||||
QY 2221 GCATGCTCCAACTCGACACAGCAAGTGAATTAATGCTGTAACGCTGCTGCGACCTTC 2280
DB |||||
QY 2281 AAGGCTCAACCATCCAAACCGTTTCTGCTGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2340
DB |||||
QY 2281 AAGGCTCAACCATCCAAACCGTTTCTGCTGCTGCTGAGCGCAGGTATTAACGCGCCAGT 2340
DB |||||
QY 2341 CCGAACAAAGCTGGCAAGAGTCTCTCGAAAACTATCTGCTGATGTAAGGTCTG 2400
DB |||||
QY 2341 CCGAACAAAGCTGGCAAGAGTCTCTCGAAAACTATCTGCTGATGTAAGGTCTG 2400
DB |||||
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGCTGCTGCTGCTGAGCTGAACTTTACGAGGAAG 2460
DB |||||
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGCTGCTGCTGCTGAGCTGAACTTTACGAGGAAG 2460
DB |||||
QY 2461 TTGCGGAAAGATCCAGTATTGCGGCAACCATGGAACACCGCCAGAAAGGTGAATCATG 2520
DB |||||
QY 2461 TTGCGGAAAGATCCAGTATTGCGGCAACCATGGAACACCGCCAGAAAGGTGAATCATG 2520
DB |||||
QY 2521 CCGAACATCCCGCAGATGCTCGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
DB |||||
QY 2521 CCGAACATCCCGCAGATGCTCGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
DB |||||
QY 2581 GCCAGCGCTGCTGAGACTGCTGATGAAGCCCTGAAAGCGCGAGACTTAATTCGAGCTCG 2640
DB |||||
QY 2581 GCCAGCGCTGCTGAGACTGCTGATGAAGCCCTGAAAGCGCGAGACTTAATTCGAGCTCG 2640
DB |||||
QY 2641 AACAAACAAACAAATTAACAACTCGGGATCGAGGGAAGGATTTTTCAGAAATTC 2700
DB |||||
QY 2641 AACAAACAAACAAATTAACAACTCGGGATCGAGGGAAGGATTTTTCAGAAATTC 2700
DB |||||

RESULT 3

US-09-265-617B-1
; Sequence 1, Application US/09265617B
; Patent No. 6372883
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R.
; APPLICANT: Hurst, David N.
; APPLICANT: Jones, Philip S.
; APPLICANT: Kay, Paul B.
; APPLICANT: Raynham, Tony M.
; APPLICANT: Wilson, Francis X.
; TITLE OF INVENTION: Antiviral Medicaments
; FILE REFERENCE: 20052 antiviral medicaments
; CURRENT APPLICATION NUMBER: US/09/265,617B
; CURRENT FILING DATE: 1999-03-10
; PRIOR APPLICATION NUMBER: GB9806815.8
; PRIOR FILING DATE: 1998-03-30
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1

; LENGTH: 7475									
; TYPE: DNA									
; ORGANISM: Artificial Sequence									
; FEATURE:									
; OTHER INFORMATION: Description of Artificial Sequence: Vector and									
; OTHER INFORMATION: Gene Fragments									
US-09-265-617B-1									
Query Match 81.8%; Score 2700; DB 3; Length 7475;									
Beat Local Similarity 100.0%; Pred. No. 0;									
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	CGGACCAACATCAATGGTGTCAAAACCTTTCGGGGTATGGCATGATAGCCCGGAAGAGA	60						
DB	1	CGGACCAACATCAATGGTGTCAAAACCTTTCGGGGTATGGCATGATAGCCCGGAAGAGA	60						
QY	61	GTCAATTCAGGGTGTGAATGTGAACACAGTAACGTTATACGATGTCGAGAGTATGCGG	120						
DB	61	GTCAATTCAGGGTGTGAATGTGAACACGTAACGTTATACGATGTCGAGAGTATGCGG	120						
QY	121	GTGTCCTTTATCAGACCGTTTCCCGGTGTGAACACAGGCCAGCCAGCGTTTCTGCGAAAA	180						
DB	121	GTGTCCTTTATCAGACCGTTTCCCGGTGTGAACACAGGCCAGCCAGCGTTTCTGCGAAAA	180						
QY	181	CGCGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTTACATTCCTCAACCCGCGTGGCAC	240						
DB	181	CGCGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTTACATTCCTCAACCCGCGTGGCAC	240						
QY	241	AACAACTGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGGCCACTCCAGTCTGGGCCCTGC	300						
DB	241	AACAACTGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGGCCACTCCAGTCTGGGCCCTGC	300						
QY	301	ACGCGCCGTGCAAAATTTGTGCGGGGATTAATCTCGCGCCGATCAACTGGGTGCGCAGCG	360						
DB	301	ACGCGCCGTGCAAAATTTGTGCGGGGATTAATCTCGCGCCGATCAACTGGGTGCGCAGCG	360						
QY	361	TGGTGGTGTGATGTGTAGAACGAAGCGCGTGAAGCCGTGAAGCGGCGGTGCAATTC	420						
DB	361	TGGTGGTGTGATGTGTAGAACGAAGCGCGTGAAGCCGTGAAGCGGCGGTGCAATTC	420						
QY	421	TTCTCGCGCAACGCGTCAGTGGGCTGATCATTTAACTATCCGCTGGATGACAGGATGCCA	480						
DB	421	TTCTCGCGCAACGCGTCAGTGGGCTGATCATTTAACTATCCGCTGGATGACAGGATGCCA	480						
QY	481	TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA	540						
DB	481	TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA	540						
QY	541	CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600						
DB	541	CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600						
QY	601	TGGTCGATGGGTGTCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTCTGCTCGG	660						
DB	601	TGGTCGATGGGTGTCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTCTGCTCGG	660						
QY	661	CGCGTCTGCGTCTGGCTGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
DB	661	CGCGTCTGCGTCTGGCTGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
QY	721	CGGAACGGGAAGGCGACTGGAGTGCGATGTCGGGTTTCAACAAACCATGTCAAAATGCTGA	780						
DB	721	CGGAACGGGAAGGCGACTGGAGTGCGATGTCGGGTTTCAACAAACCATGTCAAAATGCTGA	780						
QY	781	ATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGGCAA	840						
DB	781	ATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGGCAA	840						
QY	841	TGCGGCGCAATTAACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG	900						
DB	841	TGCGGCGCAATTAACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATAG	900						
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCATCAAAACAGGATTTTC	960						

DB	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCATCAAAACAGGATTTTC	960						
QY	961	GCCTGCTGGGGCAAAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGGCGGTGA	1020						
DB	961	GCCTGCTGGGGCAAAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGGCGGTGA	1020						
QY	1021	AGGGCAATCAGCTGTGTCCCGTCTCATCTGGTGAAGAAAAACCAACCTGGCGGCCAATA	1080						
DB	1021	AGGGCAATCAGCTGTGTCCCGTCTCATCTGGTGAAGAAAAACCAACCTGGCGGCCAATA	1080						
QY	1081	CGCAAAACCGCTCTCCCGCGGTTGGCCGATTCATTAATGAGCTGCGACGACAGGTTT	1140						
DB	1081	CGCAAAACCGCTCTCTCCCGCGGTTGGCCGATTCATTAATGAGCTGCGACGACAGGTTT	1140						
QY	1141	CCGACTCGAAAGCGGGCAGTGAGCGCAACCACTAATTAATGTGAGTTAGCTCACTCATTTAG	1200						
DB	1141	CCGACTCGAAAGCGGGCAGTGAGCGCAACCACTAATTAATGTGAGTTAGCTCACTCATTTAG	1200						
QY	1201	GCACAAATTCATGTTTGACAGCTTATCATGCACTGCAAGGTCACCAATGCTTCTGGCG	1260						
DB	1201	GCACAAATTCATGTTTGACAGCTTATCATGCACTGCAAGGTCACCAATGCTTCTGGCG	1260						
QY	1261	TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTCTGTAATCACTGCATAAATCG	1320						
DB	1261	TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTCTGTAATCACTGCATAAATCG	1320						
QY	1321	TGTGCTCAAGCGCACCTCCCGTCTTGATTAATGTTTTTGGCGCGACATCATACGGTT	1380						
DB	1321	TGTGCTCAAGCGCACCTCCCGTCTTGATTAATGTTTTTGGCGCGACATCATACGGTT	1380						
QY	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTATTAATGTGTGA	1440						
DB	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTATTAATGTGTGA	1440						
QY	1441	ATTGTGAGCGGATTAACAAATTTTCAACAGGAACACGAGTCGGTTAGGTGTTTTTCA	1500						
DB	1441	ATTGTGAGCGGATTAACAAATTTTCAACAGGAACACGAGTCGGTTAGGTGTTTTTCA	1500						
QY	1501	GCACCTTCAACCAAGGACCATAGATTATGAAGAACTGAAGAAAGGTAATCTGTAATCTGG	1560						
DB	1501	GCACCTTCAACCAAGGACCATAGATTATGAAGAACTGAAGAAAGGTAATCTGTAATCTGG	1560						
QY	1561	ATTAACCGCGATTAAGGCTATAACCGGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT	1620						
DB	1561	ATTAACCGCGATTAAGGCTATAACCGGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT	1620						
QY	1621	ACCGGAATTAAGTACACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTTCCACAGGT	1680						
DB	1621	ACCGGAATTAAGTACACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTTCCACAGGT	1680						
QY	1681	GGCGCAACTGGCGATGGCGCTGACATTAATCTTCTGGGCAACAGCCGCTTGGTGGCTAC	1740						
DB	1681	GGCGCAACTGGCGATGGCGCTGACATTAATCTTCTGGGCAACAGCCGCTTGGTGGCTAC	1740						
QY	1741	GCTCAATCTGGCTGTGTGGCTGAAATCAACCCGGACAAAGCGTTCCAGACAAAGCTGAT	1800						
DB	1741	GCTCAATCTGGCTGTGTGGCTGAAATCAACCCGGACAAAGCGTTCCAGACAAAGCTGAT	1800						
QY	1801	CGGTTTACCTGGGATGGCGTACGTTTACACCGCAAGCTGATTTGCTTACCCGATCCGTT	1860						
DB	1801	CGGTTTACCTGGGATGGCGTACGTTTACACCGCAAGCTGATTTGCTTACCCGATCCGTT	1860						
QY	1861	GAAGGTTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAACCCGCCCAAAACCTGGGAA	1920						
DB	1861	GAAGGTTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAACCCGCCCAAAACCTGGGAA	1920						
QY	1921	GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGGTAAGAGCGCTGATGTTCAAC	1980						
DB	1921	GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGGTAAGAGCGCTGATGTTCAAC	1980						
QY	1981	CTGCAAGAACCGTACTTTCACCTGGCGCTGATTTGCTGCTGACGGGGGTTATCGGTTCAAG	2040						

Db 1981 CTGCAAGAACCGTACTTCACTGGCGCGTGATTGCTGTGACGCGGGGTTATCGGTTCAAG 2040
QY 2041 TATGAAAAACGCGCAAGTACGACATTAAGACGTGGCGTGGATAAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGCGCAAGTACGACATTAAGACGTGGCGTGGATAAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGTTGACCTGATTAATAAACAACACATGAATGCGACACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTGACCTGATTAATAAACAACACATGAATGCGACACACCGATTAC 2160
QY 2161 TCCATCGCAAGAGCTGCCCTTTAATAAGCGCAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAAGAGCTGCCCTTTAATAAGCGCAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCAACATCGACACACAGCAAGTGAATTAATGTTAAACGCTACTGCCGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACACAGCAAGTGAATTAATGTTAAACGCTACTGCCGACCTTC 2280
QY 2281 AAGGTCACACCATCAACACCGTTCTGTTGGCGTGTGAGCGAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTCACACCATCAACACCGTTCTGTTGGCGTGTGAGCGAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAATACTATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAATACTATCTGCTGACTGATGAAGGTCTG 2400
QY 2401 GAAGCGGTAAATAAGACAAACCGCTGGTGGCGTGTGAGCGTGAAGTCTTAACGAGGAAG 2460
Db 2401 GAAGCGGTAAATAAGACAAACCGCTGGTGGCGTGTGAGCGTGAAGTCTTAACGAGGAAG 2460
QY 2461 TTGGCGAAGATCCACGTTATGCGCCCAACATGGAAGCGCCGAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGTTATGCGCCCAACATGGAAGCGCCGAGAAAGGTGAATCATG 2520
QY 2521 CCGAACAATCCCGCAGATGTCGCTTCTGTTGATGCGTGGCTGCTGAGTCTTAACGAGGT 2580
Db 2521 CCGAACAATCCCGCAGATGTCGCTTCTGTTGATGCGTGGCTGCTGAGTCTTAACGAGGT 2580
QY 2581 GCCAGCGTGTGAGTGTGATGAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTGTGAGTGTGATGAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACACACAAACATAACATAACAAACAAACCTCGGATCGAGGAGGATTCAGAAATTC 2700
Db 2641 AACACACAAACATAACATAACAAACAAACCTCGGATCGAGGAGGATTCAGAAATTC 2700

RESULT 4

US-09-027-169-3/c
; Sequence 3, Application US/09027169
; Patent No. 6420524
; GENERAL INFORMATION:
; APPLICANT: CRAIG, NANCY L
; TITLE OF INVENTION: GAIN OF FUNCTION MUTATIONS IN
; TITLE OF INVENTION: ATP-DEPENDENT TRANSCRIPTION PROTEINS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Anne Brown (Alston & Bird, LLP)
; STREET: 3605 Glenwood Ave.
; CITY: Raleigh
; STATE: NC
; COUNTRY: USA
; ZIP: 27608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: US/09/027,169
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:

; NAME: Brown, Anne
; REGISTRATION NUMBER: 36,463
; REFERENCE/DOCKET NUMBER: 5789-3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919 420 2205
; TELEFAX: 919 881 3175
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5926 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "pEM delta R.adj to 1"
; US-09-027-169-3

Query Match 43.0%; Score 1417.6; DB 3; Length 5926;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 1455; Conservative 0; Mismatches 9; Indels 12; Gaps 2;

QY 3 GACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGCGCATGATAGCCCGGAGAGAGT 62
Db 1564 GACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGCGCATGATAGCCCGGAGAGAGT 1505
QY 63 CAATTCAAGGGTGGTGAATGTGAACACAGTAAACGTTATACGATGTCGAGAGTATGCCGGT 122
Db 1504 CAATTCAAGGGTGGTGAATGTGAACACAGTAAACGTTATACGATGTCGAGAGTATGCCGGT 1445
QY 123 GTCTCTTATCAGACCGTTTCCCGGTGGTGAACACAGCCAGCCACGTTTCTGCGAAAAACG 182
Db 1444 GTCTCTTATCAGACCGTTTCCCGGTGGTGAACACAGCCAGCCACGTTTCTGCGAAAAACG 1385
QY 183 CGGGAAGAGTGGAGCGGCGATGGCGGAGCTGAATTACATTCCTCAACCGCTGGACAA 242
Db 1384 CGGGAAGAGTGGAGCGGCGATGGCGGAGCTGAATTACATTCCTCAACCGCTGGACAA 1325
QY 243 CAATCGCGCGCAACACAGTGTGCTGATTGGCGTTGCGACCTCCAGTCTGGGCCCTGCAC 302
Db 1324 CAATCGCGCGCAACACAGTGTGCTGATTGGCGTTGCGACCTCCAGTCTGGGCCCTGCAC 1265
QY 303 GCSCGCTCGCAAAATTGTCGCGCGATTAATCTCGCGCGGATCAACTGGGTGCAGCGTG 362
Db 1264 GCSCGCTCGCAAAATTGTCGCGCGATTAATCTCGCGCGGATCAACTGGGTGCAGCGTG 1205
QY 363 GTGCTGTGCAATGTAGAAAGCGCGTTCGAGAGCTGTAAAGCGCGGTGCAAACTCTT 422
Db 1204 GTGCTGTGCAATGTAGAAAGCGCGTTCGAGAGCTGTAAAGCGCGGTGCAAACTCTT 1145
QY 423 CTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACCGAGTGCATTT 482
Db 1144 CTCGCGCAACGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACCGAGTGCATTT 1085
QY 483 GCTGTGGAAGCTGCTGCACTAATGTTCCGGGCTTAATTTCTGATGCTCTGACCAAGACA 542
Db 1084 GCTGTGGAAGCTGCTGCACTAATGTTCCGGGCTTAATTTCTGATGCTCTGACCAAGACA 1025
QY 543 CCATCAACAGATTAATTTCTCCCATGAAGCGGTACGCGATCGGCGTGGAGCATCTG 602
Db 1024 CCATCAACAGATTAATTTCTCCCATGAAGCGGTACGCGATCGGCGTGGAGCATCTG 965
QY 603 GTGCAATTGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCCTGCTCGGG 662
Db 964 GTGCAATTGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCCTGCTCGGG 905
QY 663 CGTCTCGGCTGCTGCGCTGGCATTAATCTCACTCGCAATCAAAATTCAGCGGATAGCG 722
Db 904 CGTCTCGGCTGCTGCGCTGGCATTAATCTCACTCGCAATCAAAATTCAGCGGATAGCG 845
QY 723 GAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTTCAACAAACCATGCAAAATGCTGAAT 782
Db 844 GAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTTCAACAAACCATGCAAAATGCTGAAT 785
QY 783 GAGGCGCATGTTCCCACTCGGATGCTGGTTGCGCAACGATCAGATGGCGCTGGGCGCAATG 842

|||||
784 GAGGCGATCGTTCCCACTGCGATGCTGGTGGCAACGATCAGATGGCGCTGGGCGCAATG 725
Qy CGCGCATTTACCGAGTCCGGGCTGGCGCTGGGTGGGATATCTCGGTAGTGGGATACGAC 902
Db CGCGCATTTACCGAGTCCGGGCTGGCGCTGGGTGGGATATCTCGGTAGTGGGATACGAC 665
Qy GATACCGAAGACAGCTCATGTTATATCCGGCGTTAAACCATCAACAGGATTTTCG 962
Db GATACCGAAGACAGCTCATGTTATATCCGGCGTTAAACCATCAACAGGATTTTCG 605
Qy CTGCTGGGCGAAACAGAGCTGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAG 1022
Db CTGCTGGGCGAAACAGAGCTGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGAAG 545
Qy GGCATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAGAAACACACCTCGCGCCCAATACG 1082
Db GGCATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAGAAACACACCTCGCGCCCAATACG 485
Qy CAAACCGCTCTCCCGCGCTGCGGATTCATTAATGCGAGCTGGCAGCAGAGTTTCC 1142
Db CAAACCGCTCTCCCGCGCTGCGGATTCATTAATGCGAGCTGGCAGCAGAGTTTCC 425
Qy CGACTGGAAGCGGCGAGTGAAGCAACGCAATTAATGTGAGTTAGCTCACTATTAGG 1202
Db CGACTGGAAGCGGCGAGTGAAGCAACGCAATTAATGTGAGTTAGCTCACTATTAGG 376
Qy ACAATTTCTCATGTTTGACAGCTTATCATCGACTGCACGCTGACCAATGCTTCTGGCGTC 1262
Db GAATTTGATCTGGTTTGACAGCTTATCATCGACTGCACGCTGACCAATGCTTCTGGCGTC 316
Qy AGGAGCCATCGGAAGCTGTGGTATGGCTGTGCGAGTGGTAAATCACTGCATTAATTCGTG 1322
Db AGGAGCCATCGGAAGCTGTGGTATGGCTGTGCGAGTGGTAAATCACTGCATTAATTCGTG 256
Qy TCGCTCAAGGCGCACTCCGCTCGGATAGTGTGCGCGGACATCAACCGTTCT 1382
Db TCGCTCAAGGCGCACTCCGCTCGGATAGTGTGCGCGGACATCAACCGTTCT 196
Qy GGCATATTTCTGAATGAGCTGTTGACAAATTAATCAT- CGGCTCGTATATGTGTGAA 1441
Db GGCATATTTCTGAATGAGCTGTTGACAAATTAATCATTAATCATCCGGCTGTATATGTGAA 136
Qy TTGTGAGCGGATPAACAAATTTACACAGGAAACAGCC 1477
Db TTGTGAGCGGATPAACAAATTTACACAGGAAACAGAC 100

RESULT 5

US-09-640-882-2
; Sequence 2, Application US/09640882
; Patent No. 6720142
; GENERAL INFORMATION:
; APPLICANT: Hall, Barry G.
; TITLE OF INVENTION: METHOD OF DETERMINING EVOLUTIONARY POTENTIAL OF MUTANT
; TITLE OF INVENTION: RESISTANCE GENES AND USE THEREOF TO SCREEN FOR DRUG
; TITLE OF INVENTION: EFFICACY
; FILE REFERENCE: 176/60851
; CURRENT APPLICATION NUMBER: US/09/640,882
; CURRENT FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/149,813
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 5201
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid
; OTHER INFORMATION: pACSE2
US-09-640-882-2

Query Match 41.3%; Score 1362; DB 4; Length 5201;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1380; Conservative 0; Mismatches 10; Indels 2; Gaps 1;
Qy 3 GACACCAATCGAATGGTGC AAAA CCTTTTCGCGGTATGGCATATAGCGCCCGGAGAGAGT 62
Db 866 GACACCAATCGAATGGTGC AAAA CCTTTTCGCGGTATGGCATATAGCGCCCGGAGAGAGT 925
Qy 63 CAATTCAGGGTGGTGAATGTGA AACCAAGTAAAGTTATACGATTCGCGAGAGATGCGCGT 122
Db 926 CAATTCAGGGTGGTGAATGTGA AACCAAGTAAAGTTATACGATTCGCGAGAGATGCGCGT 985
Qy 123 GTCTCTTATCAGACCGTTTCGCGGTGGTGAACAGCGCCAGCGCTTCTCGGAAAACG 182
Db 986 GTCTCTTATCAGACCGTTTCGCGGTGGTGAACAGCGCCAGCGCTTCTCGGAAAACG 1045
Qy 183 CGGAAAAGTGGAAAGCGGCGATCGCGAGCTGAAATTAATTCATTCCTCAACCGCGTGCAAA 242
Db 1046 CGGAAAAGTGGAAAGCGGCGATCGCGAGCTGAAATTAATTCATTCCTCAACCGCGTGCAAA 1105
Qy 243 CAATCGCGGCAACAGTCTGTTGCTGATTCGCGGTGGCCACTCGAGTCTGGCCCTGCAC 302
Db 1106 CAATCGCGGCAACAGTCTGTTGCTGATTCGCGGTGGCCACTCGAGTCTGGCCCTGCAC 1165
Qy 303 CGCGCTGCGCAAAATGTTCGCGCGATTAATTCCTCGCGCGATCACTGGGTGCCAGCGTG 362
Db 1166 CGCGCTGCGCAAAATGTTCGCGCGATTAATTCCTCGCGCGATCACTGGGTGCCAGCGTG 1225
Qy 363 GTGGTGTGATGTTAGAACGAGCGGCTCGAAGCTCTGAAGCGGCGGTGCACAACTCT 422
Db 1226 GTGGTGTGATGTTAGAACGAGCGGCTCGAAGCTCTGAAGCGGCGGTGCACAACTCT 1285
Qy 423 CTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCACTATCCGCTGGATGACAGGATGCCATT 482
Db 1286 CTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCACTATCCGCTGGATGACAGGATGCCATT 1345
Qy 483 CTTGTGGAAGCTGCTGCACATAATGTTCCGCGGTATTTCTTGTATCTCTGACAGACA 542
Db 1346 CTTGTGGAAGCTGCTGCACATAATGTTCCGCGGTATTTCTTGTATCTCTGACAGACA 1405
Qy 543 CCCATCAACAGTATTTATTTCTCCCATGAAGAGGTACGCGACTGGGCGTGGAGCATCTG 602
Db 1406 CCCATCAACAGTATTTATTTCTCCCATGAAGAGGTACGCGACTGGGCGTGGAGCATCTG 1465
Qy 603 CTCGCATTGGGTCAACAGCAAAATCGCGCTGTTAGCGGCGCCATTAAGTTCTGTCTCGCG 662
Db 1466 CTCGCATTGGGTCAACAGCAAAATCGCGCTGTTAGCGGCGCCATTAAGTTCTGTCTCGCG 1525
Qy 663 CGTCTGCGTCTGGCTGGCGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
Db 1526 CGTCTGCGTCTGGCTGGCGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 1585
Qy 723 GAACGGGAAGCGACTGGAGTGCCATGTCGCGTTTCAACAAACCATCAAAATGCTGAAT 782
Db 1586 GAACGGGAAGCGACTGGAGTGCCATGTCGCGTTTCAACAAACCATCAAAATGCTGAAT 1645
Qy 783 GAGGCGATCTGTTCCCATCTGCGATGCTGGTTGCGCAACGATCAGATGCGCGTGGGCGCAATG 842
Db 1646 GAGGCGATCTGTTCCCATCTGCGATGCTGGTTGCGCAACGATCAGATGCGCGTGGGCGCAATG 1705
Qy 843 CGCGCATTTACCGAGTCCGGGCTGGCGTGGTGGGATATCTCGGTAGTGGGATACGAC 902
Db 1706 CGCGCATTTACCGAGTCCGGGCTGGCGTGGTGGGATATCTCGGTAGTGGGATACGAC 1765
Qy 903 GATACCGAAGACAGCTCATGTTATATCCGCGGTTAACCAACCATCAAAACGATTTTCG 962
Db 1766 GATACCGAAGACAGCTCATGTTATATCCGCGGTCAACCAACCATCAAAACGATTTTCG 1825
Qy 963 CTGCTGGGCGCAACAGCGGTGGACCGCTTGTGCAACTCTCTCAGGGCGCAGCGGTGAAG 1022
Db 1826 CTGCTGGGCGCAACAGCGGTGGACCGCTTGTGCAACTCTCTCAGGGCGCAGCGGTGAAG 1885
Qy 1023 GGCATCAGCTGTTGCGCGTCTCACTGTTGAAAGAAACCAACCTGGCGCCCAATACG 1082

1886 GGCNAATCAGTGTGTGCGCTCTCACTGGTGAAGAAAAACACCTCGCGCCCAATAGC 1945
1083 CAAACCGCTCTCCCGCGCGTGGCGGATTCATTAATGACGTGGCAGCAGAGTTTCC 1142
1946 CAAACCGCTCTCCCGCGCGTGGCGGATTCATTAATGACGTGGCAGCAGAGTTTCC 2005
1143 CGACTGGAAGGGGCGAGTGAGCGCAACGCAATTAATGTAGCTAGCTCACTCAATTAGC 1202
2006 CGACTGGAAGGGGCGAGTGAGCGCAACGCAATTAATGTAGCTAGCTCACTCAATTAGC 2063
1203 ACAATTCATGTTGACAGCTTATCATGCACTGACGCGTGACCAATGCTTCTGCGGTC 1262
2064 TGAATTCATGTTGACAGCTTATCATGCACTGACGCGTGACCAATGCTTCTGCGGTC 2123
1263 AGGCAGCCATCGAAGCTGTGGTATGGCTGTGCAAGTGTAAATCACTGCATTAATTCGTG 1322
2124 AGGCAGCCATCGAAGCTGTGGTATGGCTGTGCAAGTGTAAATCACTGCATTAATTCGTG 2183
1323 TCGCTCAAGGCGCACTCCCGTTCTCGATAATGTTTTTGGCGCGACATCAATCGGTTCT 1382
2184 TCGCTCAAGGCGCACTCCCGTTCTCGATAATGTTTTTGGCGCGACATCAATCGGTTCT 2243
1383 GGCNAATATCT 1394
2244 GGCNAATATCT 2255

RESULT 6
US-09-640-882-3
; Sequence 3, Application US/09640882
; Patent No. 6720142
; GENERAL INFORMATION:
; APPLICANT: Hall, Barry G.
; TITLE OF INVENTION: METHOD OF DETERMINING EVOLUTIONARY POTENTIAL OF MUTANT
; TITLE OF INVENTION: RESISTANCE GENES AND USE THEREOF TO SCREEN FOR DRUG
; TITLE OF INVENTION: EFFICACY
; FILE REFERENCE: 176/60851
; CURRENT APPLICATION NUMBER: US/09/640,882
; CURRENT FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/149,813
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3
; LENGTH: 5201
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid
; OTHER INFORMATION: PACSE3
US-09-640-882-3

Query Match 41.3%; Score 1362; DB 4; Length 5201;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1380; Conservative 0; Mismatches 10; Indels 2; Gaps 1;
3 GACACCATCGAATCGTGCAAAACCTTTTCGCGGTATGCGCATGATAGCGCCCGGAAGAGAGT 62
866 GACACCATCGAATCGTGCAAAACCTTTTCGCGGTATGCGCATGATAGCGCCCGGAAGAGAGT 925
63 CAATTACAGGTGGTGAATGTGAACCCAGTAACTGTTATACGATGCGCAGAGTATGCCGTT 122
926 CAATTACAGGTGGTGAATGTGAACCCAGTAACTGTTATACGATGCGCAGAGTATGCCGTT 985
123 GTCTCTTATCAGACCGTTTCCCGCGTGTGAACCCAGGCGCACCGCTTCTCGGAAACG 182
986 GTCTCTTATCAGACCGTTTCCCGCGTGTGAACCCAGGCGCACCGCTTCTCGGAAACG 1045
183 CGGGAAGAGTGGNAGCGCGGATGCGGAGCTGAATTCATTCCTCCACCGCGTGGCACA 242
1046 CGGGAAGAGTGGNAGCGCGGATGCGGAGCTGAATTCATTCCTCCACCGCGTGGCACA 1105

243 CAACCTGGCGGCAACAGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGCGAC 302
1106 CAACCTGGCGGCAACAGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGCGAC 1165
303 GCGCGCTGCGCAAAATGTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGGCGT 362
1166 GCGCGCTGCGCAAAATGTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGGCGT 1225
363 GTGTGTGATGTTAGTAAACGAGCGGCTGAAAGCTTGTAAAGCGGCGGTGCAAACTCTT 422
1226 GTGTGTGATGTTAGTAAACGAGCGGCTGAAAGCTTGTAAAGCGGCGGTGCAAACTCTT 1285
423 CTGCGCAACGCGTCAGTGGGCTGATCAATTAATCTATCGGCTGATGATGACGAGATGCCATT 482
1286 CTGCGCAACGCGTCAGTGGGCTGATCAATTAATCTATCGGCTGATGATGACGAGATGCCATT 1345
483 GCTGTGGAAGCTGCTGCATCAATGTTCCGGCGTTATTTCTGATGTTCTCTGACGACACA 542
1346 GCTGTGGAAGCTGCTGCATCAATGTTCCGGCGTTATTTCTGATGTTCTCTGACGACACA 1405
543 CCNATCAACAGTATTTTCTCCATGAAGAGCGGTACGCGACTGGGCGTGGAGCATCTG 602
1406 CCCATCAACAGTATTTTCTCCATGAAGAGCGGTACGCGACTGGGCGTGGAGCATCTG 1465
603 GTGCAATTGGGTCAACAGCAAAATCGGCTGTAGCGGCGCATTTAGTCTGTTCTGCGCG 662
1466 GTGCAATTGGGTCAACAGCAAAATCGGCTGTAGCGGCGCATTTAGTCTGTTCTGCGCG 1525
663 CGTCTGCGTCTGCTGGCTGGCATAAATATCTCACTGCAATCAAAATTCAGCGCATAGCG 722
1526 CGTCTGCGTCTGCTGGCTGGCATAAATATCTCACTGCAATCAAAATTCAGCGCATAGCG 1585
723 GAACGGGAAGGCACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGAAT 782
1586 GAACGGGAAGGCACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGAAT 1645
783 GAGGCGATCTTCCCACTCGGATGCTGTTGCCAAGATCAGATGCGCTGGGCGCAATG 842
1646 GAGGCGATCTTCCCACTCGGATGCTGTTGCCAAGATCAGATGCGCTGGGCGCAATG 1705
843 GCGGCCATTACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACGAC 902
1706 GCGGCCATTACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACGAC 1765
903 GATACCGGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAAAAGAGATTTTCG 962
1766 GATACCGGAAGACAGCTCATGTTATATCCCGCGTTAACCCACCATCAAAAGAGATTTTCG 1825
963 CTGCTGGGCAACACGAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCGAGCGGTGAAG 1022
1826 CTGCTGGGCAACACGAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCGAGCGGTGAAG 1885
1023 GGCNAATCAGTGTGCGCGTCTCACTGGTGAAGAAAAACCACTCGCGCGCCCAATAGC 1082
1886 GGCNAATCAGTGTGCGCGTCTCACTGGTGAAGAAAAACCACTCGCGCGCCCAATAGC 1945
1083 CAAACCGCTCTCCCGCGCGTGGCGGATTCATTAATGACGTGGCAGCAGAGTTTCC 1142
1946 CAAACCGCTCTCCCGCGCGTGGCGGATTCATTAATGACGTGGCAGCAGAGTTTCC 2005
1143 CGACTGGAAGGGGCGAGTGAGCGCAACGCAATTAATGTAGCTAGCTCACTCAATTAGC 1202
2006 CGACTGGAAGGGGCGAGTGAGCGCAACGCAATTAATGTAGCTAGCTCACTCAATTAGC 2063
1203 ACAATTCATGTTGACAGCTTATCATGCACTGACGCGTGACCAATGCTTCTGCGGTC 1262
2064 TGAATTCATGTTGACAGCTTATCATGCACTGACGCGTGACCAATGCTTCTGCGGTC 2123
1263 AGGCAGCCATCGAAGCTGTGGTATGGCTGTGCAAGTGTAAATCACTGCATTAATTCGTG 1322
2124 AGGCAGCCATCGAAGCTGTGGTATGGCTGTGCAAGTGTAAATCACTGCATTAATTCGTG 2183
1323 TCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAATCGGTTCT 1382

Db	2184	TCGCTCAAGGCGGCACTCCCGTTCTGGATAATGTTTTTGGCGGACATCATAACGGTTCT	2243
Qy	1383	GGCAAAATATTCT 1394	
Db	2244	GGCAAAATATTCT 2255	
RESULT 7			
US-08-778-717-5			
; Sequence 5, Application US/08778717			
; Patent No. 6603689			
; GENERAL INFORMATION:			
; APPLICANT: UENO, EIICHI			
; APPLICANT: NOBUYUKI, FUJII			
; APPLICANT: OKADA, MAGAHISA			
; TITLE OF INVENTION: FUSED DNA SEQUENCE, FUSED PROTEIN			
; TITLE OF INVENTION: EXPRESSED FROM SAID FUSED DNA SEQUENCE AND METHOD FOR			
; TITLE OF INVENTION: EXPRESSING SAID DNA SEQUENCE			
; NUMBER OF SEQUENCES: 21			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,			
; ADDRESSEE: P.C.			
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400			
; CITY: ARLINGTON			
; STATE: VA			
; COUNTRY: USA			
; ZIP: 22202			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/778,717			
; FILING DATE: 12-DEC-1996			
; CLASSIFICATION: 536			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: JP 352225/1995			
; FILING DATE: 28-DEC-1995			
; ATTORNEY/AGENT INFORMATION:			
; NAME: OBLON, NORMAN F.			
; REGISTRATION NUMBER: 24,618			
; REFERENCE/DOCKET NUMBER: 2084-031-0			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 703-413-3000			
; TELEFAX: 703-413-2220			
; INFORMATION FOR SEQ ID NO: 5:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 4557 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: double			
; TOPOLOGY: circular			
; MOLECULE TYPE: other nucleic acid			
; ORIGINAL SOURCE:			
; ORGANISM: E. COLI			
; STRAIN: BL21 (DE3)			
; PUBLICATION INFORMATION:			
; AUTHORS: NOBUYUKI FUJII ET AL,			
; TITLE: FUSED DNA SEQUENCE, FUSED PROTEIN EXPRESSED			
; TITLE: FROM SAID FUSED DNA SEQUENCE AND METHOD OF			
; TITLE: EXPRESSING SAID FUSED PROTEIN			
; RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 TO 4557			
US-08-778-717-5			
Query Match 36.5%; Score 1204; DB 4; Length 4557;			
Best Local Similarity 100.0%; Pred.No. 0;			
Matches 1204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	CCGACACGATCGAATGGTGCAGAAACCTTTTCGCGGTATGGCATGTAGCGCCCGAAGAGA	60
Db	2563	CCGACACCATCGAATGGTGCAGAAACCTTTTCGCGGTATGGCATGTAGCGCCCGAAGAGA	2622

Db 3703 CCCGACTGGAAGCGGCGAGTGGCGCAACGCAATTAATGTAGTACTACTATTAG 3762
Qy 1201 GCAC 1204
Db 3763 GCAC 3766

RESULT 8
US-08-148-675A-2
; Sequence 2, Application US/08148675A
; Patent No. 5506121
; GENERAL INFORMATION:
; APPLICANT: Skerra, Arne; Schmidt, Thomas
; TITLE OF INVENTION: FUSION PEPTIDES WITH BINDING ACTIVITY FOR
; TITLE OF INVENTION: STREPTAVIDIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/148.675A
; FILING DATE: 3-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 42 37 113.9
; FILING DATE: 03-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Teal, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: HUBR 1041
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3832 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

US-08-148-675A-2

Query Match 36.4%; Score 1200.8; DB 1; Length 3832;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1202; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGTGCAAAACCTTCCGGTATGGCATGATAGCCCGGAAAGAGA 60
Db 3 CCGACACCATCGAATGTGCGCAAAACCTTCCGGTATGGCATGATAGCCCGGAAAGAGA 62

Qy 61 GTCAATTCAAGGTGGTGAATGTGAACAGTAACGTTATACGATGTGCGAGAGTATGCCG 120
Db 63 GTCAATTCAAGGTGGTGAATGTGAACAGTAACGTTATACGATGTGCGAGAGTATGCCG 122

Qy 121 GTGTCTCTTATCAGACCGTTTCCCGGTGTGGAACAGCCAGCCAGCCAGTTCTCTCGGAAAA 180
Db 123 GTGTCTCTTATCAGACCGTTTCCCGGTGTGGAACAGCCAGCCAGCCAGTTCTCTCGGAAAA 182

Qy 181 CGCGGGAAAAAGTGGAGCGCGATGCGGAGCTGAATTACATCCCAACCGGTGGCAC 240
Db 183 CGCGGGAAAAAGTGGAGCGCGATGCGGAGCTGAATTACATCCCAACCGGTGGCAC 242

Qy 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGGCACCTCCAGTCTGGCCCTGC 300

Db 243 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 302
Qy 301 ACGCCGCGTCCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATCAATCTGGGTGCCAGCG 360
Db 303 ACGCCGCGTCCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATCAATCTGGGTGCCAGCG 362

Qy 361 TGGTGGTGTGATGCTAGAACGAGCGCGTCCGAAGCCTGTAAAGCGCGGTGCACAAATC 420
Db 363 TGGTGGTGTGATGCTAGAACGAGCGCGTCCGAAGCCTGTAAAGCGCGGTGCACAAATC 422

Qy 421 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTACGCTGGATCACCAGGATGCCA 480
Db 423 TTCTCGCGCAACCGCTCAGTGGCTGATCAATTAATCTACGCTGGATCACCAGGATGCCA 482

Qy 481 TTGCTGTGGAGCTGCGCTGCACTAATGTTCCGGGTTATTTCTTGATGTCTCTGACCAGA 540
Db 483 TTGCTGTGGAGCTGCGCTGCACTAATGTTCCGGGTTATTTCTTGATGTCTCTGACCAGA 542

Qy 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGCTACGCGACTGGGCGTGGAGCATC 600
Db 543 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGCTACGCGACTGGGCGTGGAGCATC 602

Qy 601 TGGTCGCAATTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Db 603 TGGTCGCAATTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG 662

Qy 661 CGGCTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 663 CGGCTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 722

Qy 721 CGGAACGGGAAGCGACTGGAGTGCCATGTCGGTTCCTCAACAAACCATCAAAATGCTGA 780
Db 723 CGGAACGGGAAGCGACTGGAGTGCCATGTCGGTTCCTCAACAAACCATCAAAATGCTGA 782

Qy 781 ATGAGGCGATCGTTCCCATGTCGATGTCGCAATGCCAAACGATCAGATGGCGCTGGCGCAA 840
Db 783 ATGAGGCGATCGTTCCCATGTCGATGTCGCAATGCCAAACGATCAGATGGCGCTGGCGCAA 842

Qy 841 TGGCGGCCATTACCGAGTCCGGGCTGGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 843 TGGCGGCCATTACCGAGTCCGGGCTGGGTTGGTGGGATATCTCGGTAGTGGGATACG 902

Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAAACCAACATCAACAGGATTTTC 960
Db 903 ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAAACCAACATCAACAGGATTTTC 962

Qy 961 GCCTGCTGGGGCAAAACGAGTGGGACCGCTTGTGTCGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 963 GCCTGCTGGGGCAAAACGAGTGGGACCGCTTGTGTCGCAACTCTCTCAGGGCCAGCGGTGA 1022

Qy 1021 AGGGCAATCAGCTGTTGCGCGCTCTCACTGGTGAAGAAAGAAACCCCTGGCGCCCAATA 1080
Db 1023 AGGGCAATCAGCTGTTGCGCGCTCTCACTGGTGAAGAAAGAAACCCCTGGCGCCCAATA 1082

Qy 1081 CGCAAAACCGCTCTCCCGCGCGTTGGCCGATTCATTAATGCGAGCTGGCAGCAGAGTTT 1140
Db 1083 CGCAAAACCGCTCTCCCGCGCGTTGGCCGATTCATTAATGCGAGCTGGCAGCAGAGTTT 1142

Qy 1141 CCGGACTGGAAGCGGGCAGTGAAGCGCAACGCAATTAATGCGAGTTAGTCTACTCATTTAG 1200
Db 1143 CCGGACTGGAAGCGGGCAGTGAAGCGCAACGCAATTAATGCGAGTTAGTCTACTCATTTAG 1202

Qy 1201 GCAC 1204
Db 1203 GCAC 1206

RESULT 9
US-08-487-283A-18/c
; Sequence 18, Application US/08487283A
; Patent No. 6355245
; GENERAL INFORMATION:
; APPLICANT: Evans, Mark J.

APPLICANT: Matis, Louis A.
APPLICANT: Mueller, Eileen Elliott
APPLICANT: Nye, Steven H.
APPLICANT: Rollins, Russell P.
APPLICANT: Rother, Russell P.
APPLICANT: Springhorn, Jeremy P.
APPLICANT: Squinto, Stephen P.
APPLICANT: Thomas, Thomas C.
APPLICANT: Wilkins, James A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT
OF INFLAMMATORY DISEASES
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seth A. Fidel
CITY: New Haven
STATE: Connecticut
COUNTRY: USA
ZIP: 06511
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.4Mb storage
COMPUTER: Macintosh Cetrus 610
OPERATING SYSTEM: System 7
SOFTWARE: WordPerfect 3.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,283A
FILING DATE: June 7, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/236,208
FILING DATE: 02-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Seth A. Fidel.
REGISTRATION NUMBER: 38,449
REFERENCE/DOCKET NUMBER: ALX-152.1 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203)776-1790
TELEFAX: (203)772-3655
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 5248 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Circular
MOLECULE TYPE: Other nucleic acid
DESCRIPTION: PET Trc SOS/NI
DESCRIPTION: prokaryotic expression vector
US-08-487-283A-18

Query Match 36.3%; Score 1199.2; DB 3; Length 5248;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CGGACACCATCGAATGGTGCAGAACTTTCCGGGTATGCGCATAGAGCCCGGAAGAGA 60
DB 4810 CGGACACCATCGAATGGCGCAAACTTTCCGGGTATGCGCATAGAGCCCGGAAGAGA 4751
QY 61 GTCAATTCAGGGTGTGAATGTGAACACAGTAACGTATACCATGTGCGCAGAGTATGCCG 120
DB 4750 GTCAATTCAGGGTGTGAATGTGAACACAGTAACGTATACCATGTGCGCAGAGTATGCCG 4691
QY 121 GTGTCTCTTATCAGACCGTTTCCCGGTGTGAACACAGCCAGCCACGTTTCTGCGAAAA 180
DB 4690 GTGTCTCTTATCAGACCGTTTCCCGGTGTGAACACAGCCAGCCACGTTTCTGCGAAAA 4631
QY 181 CGCGGAAAAAGTGAACGGCGATGCGGAGCTGAATTCACATCCCAACCGCGTGGCAC 240
DB 4630 CGCGGAAAAAGTGAACGGCGATGCGGAGCTGAATTCACATCCCAACCGCGTGGCAC 4571
QY 241 AACAACTGCGCGGCAAAACAGTCTGTGCTGATTTGGGGTTGCCACCTCCAGTCTGGCCCTGC 300
DB 4570 AACAACTGCGCGGCAAAACAGTCTGTGCTGATTTGGGGTTGCCACCTCCAGTCTGGCCCTGC 4511
QY 301 ACGCGCGGTGCGCAAAATGTGCGGGCGATTAAATCTCGGCGCGATCAACTGGGTGCCAGCG 360

RESULT 10

PCR-US96-05611A-21/c
; Sequence 21, Application PC/TUS9605611A
; GENERAL INFORMATION:
; APPLICANT: Mueller, John P.
; APPLICANT: Lenardo, Michael J.
; APPLICANT: McFarland, Henry F.
; APPLICANT: Matis, Louis A.
; APPLICANT: Mueller, Eileen Elliott

Db 4510 ACGCGCGTGCAGAAATGTTCGGCGGATTAATCTCGCGCGATCAACTGGGTGCCACGG 4451
QY 361 TGGTGGTTCGATGTTAGAACGAGCGCGTTCGAAGCCTGTAAAGCGCGGTGCAACAATC 420
Db 4450 TGGTGGTTCGATGTTAGAACGAGCGCGTTCGAAGCCTGTAAAGCGCGGTGCAACAATC 4391
QY 421 TTCTCTCGGCAACGGCTCAGTGGGCTGATCAATTAATATCCGCTGGATGACGAGATGCA 480
Db 4390 TTCTCGGCAACGGCTCAGTGGGCTGATCAATTAATATCCGCTGGATGACGAGATGCA 4331
QY 481 TTGTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTATTCTTCTGATGCTCTGACCCAGA 540
Db 4330 TTGTGTGGAAGCTGCCCTGCACTAATGTTCCGGCGTATTCTTCTGATGCTCTGACCCAGA 4271
QY 541 CACCCATCAACAGATATTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Db 4270 CACCCATCAACAGATATTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 4211
QY 601 TGGTCGCATTTGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db 4210 TGGTCGCATTTGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 4151
QY 661 CGCGTCTGCGCTCGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4150 CGCGTCTGCGCTCGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 4091
QY 721 CGGAACGGGAAGGCGACTGCGAGTGCATGTCGGGTTTTCAACAACCAATGCAAAATGCTGA 780
Db 4090 CGGAACGGGAAGGCGACTGCGAGTGCATGTCGGGTTTTCAACAACCAATGCAAAATGCTGA 4031
QY 781 ATGAGGGCATCGTTCCCACTCGCATGCTGGTTGCGCAACGATCAGATGCGGCTGGCGGCAA 840
Db 4030 ATGAGGGCATCGTTCCCACTCGCATGCTGGTTGCGCAACGATCAGATGCGGCTGGCGGCAA 3971
QY 841 TGGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db 3970 TGGCGCCATTACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 3911
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAAACCAACCAATCAAAAGGATTTTC 960
Db 3910 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAAACCAACCAATCAAAAGGATTTTC 3851
QY 961 GCCTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 3850 GCCTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 3791
QY 1021 AGGCAATCAGCTGTTCGCCCTCTCACTGGTGAAGAAACCAACCTGGCGGCCAATA 1080
Db 3790 AGGCAATCAGCTGTTCGCCCTCTCACTGGTGAAGAAACCAACCTGGCGGCCAATA 3731
QY 1081 CGCAACCGCTCTCTCCCGCGCTTGGCGGATTCAATTAATGCGCTGGCAGCAGAGTTT 1140
Db 3730 CGCAACCGCTCTCTCCCGCGCTTGGCGGATTCAATTAATGCGCTGGCAGCAGAGTTT 3671
QY 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACCGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Db 3670 CCCGACTGGAAGCGGGCAGTGAGCGCAACCGCAATTAATGTAGTTAGCTCACTCATTTAG 3611
QY 1201 GCAC 1204
Db 3610 GCAC 3607

```

; APPLICANT: Nye, Steven H.
; APPLICANT: Pelfrey, Clara M.
; APPLICANT: Squinto, Stephen P.
; APPLICANT: Wilkins, James A.
; TITLE OF INVENTION: Modified Myelin Protein Molecules
; NUMBER OF SEQUENCES: 299

```

```

;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: Maurice M. Klee
; STREET: 1951 Burr Street
; CITY: Fairfield
; STATE: Connecticut
; COUNTRY: USA
; ZIP: 06430
;

```

```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 0.8 Mb storage
; COMPUTER: Macintosh Centris 610
; OPERATING SYSTEM: System 7
; SOFTWARE: Microsoft Word 6.0.1
;

```

```

/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US96/05611A
/ FILING DATE: 02-MAY-1995
/ CLASSIFICATION:

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/431,644
; FILING DATE: May 2, 1995
; APPLICATION NUMBER: 08/431,648
; FILING DATE: May 2, 1995
; APPLICATION NUMBER: 08/482,114
; FILING DATE: June 7, 1995
; ATTORNEY/AGENT INFORMATION:

NAME: Klee, Maurice M.
REGISTRATION NUMBER: 30,399
REFERENCE/DOCKET NUMBER: ALX-129
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203) 255 1400
TELEFAX: (203) 254 1101

TELEFAX: (203) 234 1101
; INFORMATION FOR SEQ ID NO: 21:

SEQUENCE CHARACTERISTICS:

LENGTH: 5248 base pairs
TYPE: Nucleic Acid

STRANDEDNESS: Double

TOPOLOGY: Circular

;	MOLECULE TYPE:	Other nucleic acid
;	DESCRIPTION:	PET Trc SQ5/NT

DESCRIPTION: prokaryotic expression vector

PCT-US96-05611A-21

Query Match	36.3%	Score 1199.2	DB 5	Length 5248
Best Local Similarity	99.8%	Pred. No. 0		
Matches 1201	Conservative 0	Mismatches 3	Indels 0	Gaps 0
QY 1	CCGACACCATCGAATGGTGTGCAAAACCTTTCCGCGGTATGGCATGATACGCCCGGAACAGA	60		
DB	4810	CGGACACCATCGAATGGCGCAAAACCTTTCCGCGGTATGGCATGATGATACGCCCGGAACAGA	4751	
QY 61	GTCAATTCAGGTTGGTGAATGTGAAACCAAGTAACGTTTATACGATGTCGCAGAGTATGCCG	120		
DB	4750	GTCAATTCAGGTTGGTGAATGTGAAACCAAGTAACGTTTATACGATGTCGCAGAGTATGCCG	4691	
QY 121	GTGTCTCTTATCAGACCGTTTCCGCGTGTGTGAACCAAGCCAGCCACCGTTTCTCGGAAA	180		
DB	4690	GTGTCTCTTATCAGACCGTTTCCGCGTGTGTGAACCAAGCCAGCCACCGTTTCTCGGAAA	4631	
QY 181	CGCGGGAAAAGTGGAAACGGCGGATGCGGAGCTGAAATTAATTCCCAACCGGTGGCAC	240		
DB	4630	CGCGGGAAAAGTGGAAACGGCGGATGCGGAGCTGAAATTAATTCCCAACCGGTGGCAC	4571	
QY 241	AACAACCTGGGGGCAACACAGTCGTGCTGATTCGGCGTTGCCACTCCAGTCTGGCCCTGC	300		
DB	4570	AACAACCTGGGGGCAACACAGTCGTGCTGATTCGGCGTTGCCACTCCAGTCTGGCCCTGC	4511	
QY 301	ACGCGCGCTCGCAAAATTGTCCGGCGGATTAAATCTCGCGCCGATCAACTGGGTGCCAGCG	360		

D _b	4510	ACGCGCCGTGCGAAATTGTTCGGCGCGATTAATACTCGCGCGAGTCAACTGGGTGCACGCG	4451
Q _y	361	TGTTGGTGTCGATGTAAGAACGAAGCGCGTGCGAAGCTGTAAAGCGCGGTGCACAATC	420
D _b	4450	TGTTGGTGTCGATGTAAGAACGAAGCGCGTGCGAAGCTGTAAAGCGCGGTGCACAATC	4391
Q _y	421	TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATACTATCCGCTGGATGACCAGGATGCCA	480
D _b	4390	TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATACTATCCGCTGGATGACCAGGATGCCA	4331
Q _y	481	TTGCTGTGGGAAGCTGCCTGCACATAATGTCGGGCGTTATTTCTTGATGTCCTGACCCAGA	540
D _b	4330	TTGCTGTGGGAAGCTGCCTGCACATAATGTCGGGCGTTATTTCTTGATGTCCTGACCCAGA	4271
Q _y	541	CACCCATCAAACAGTATTAATTTCTCCCATGAAGACGGTAGCGACTGGGGGTGGAGCATC	600
D _b	4270	CACCCATCAAACAGTATTAATTTCTCCCATGAAGACGGTAGCGACTGGGGGTGGAGCATC	4211
Q _y	601	TGTTGCGCATTTGGGTACACAGCAAAATCGCGCTGTAGCGGGGCCATTAAGTCTCTGTCGG	660
D _b	4210	TGTTGCGCATTTGGGTACACAGCAAAATCGCGCTGTAGCGGGGCCATTAAGTCTCTGTCGG	4151
Q _y	661	CGCGTCTGCGTCTGGCTGGCTGGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
D _b	4150	CGCGTCTGCGTCTGGCTGGCTGGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	4091
Q _y	721	CGGAACGGGAAGCGCTGAGTGCCATGTCCGGTTTTTCAACAAAACATGCAAAATGCTGA	780
D _b	4090	CGGAACGGGAAGCGCTGAGTGCCATGTCCGGTTTTTCAACAAAACATGCAAAATGCTGA	4031
Q _y	781	ATGAGGCGCATCGTTCCCATCTGCGATGCTGTTGTCCAACAGATCAGATGGCGCTGGGGCGAA	840
D _b	4030	ATGAGGCGCATCGTTCCCATCTGCGATGCTGTTGTCCAACAGATCAGATGGCGCTGGGGCGAA	3971
Q _y	841	TGCGCGGCATTACGAGTCCGGCTGGCGTTGGTGGGATATCTCGTATGGGATACG	900
D _b	3970	TGCGCGGCATTACGAGTCCGGCTGGCGTTGGTGGGATATCTCGTATGGGATACG	3911
Q _y	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTTAAACACCATCAAAACAGATTTTC	960
D _b	3910	ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTTAAACACCATCAAAACAGATTTTC	3851
Q _y	961	GCCTGCTGGGGCAAAACGAGCTGGACCGCTTGCTGCAACTCTCTAGGGCCAGCGGTGA	1020
D _b	3850	GCCTGCTGGGGCAAAACGAGCTGGACCGCTTGCTGCAACTCTCTAGGGCCAGCGGTGA	3791
Q _y	1021	AGGGCAATCAGCTGTTGGCCGCTCTCACTGTTGAAAAAGAAAAACCCCTGGGCCCAATA	1080
D _b	3790	AGGGCAATCAGCTGTTGGCCGCTCTCACTGTTGAAAAAGAAAAACCCCTGGGCCCAATA	3731
Q _y	1081	CGCAAAACCGCTCTCCC CGCGCTTGGCGGATTCATTAAATGCAAGCTGGCACGACGTTT	1140
D _b	3730	CGCAAAACCGCTCTCCC CGCGCTTGGCGGATTCATTAAATGCAAGCTGGCACGACGTTT	3671
Q _y	1141	CCGCACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGTCTCACTCATTAG	1200
D _b	3670	CCGCACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGTCTCACTCATTAG	3611
Q _y	1201	GCAC 1204	
D _b	3610	GCAC 3607	

RESULT 11
US-10-263-103-35/c
; Sequence 35, Application US/10263103
; Patent No. 6821723
; GENERAL INFORMATION:
; APPLICANT: AVENTIS PASTEUR
; APPLICANT: Chevallier, Michel
; APPLICANT: El Habib, Raphaelle
; APPLICANT: Krell, Tino
; APPLICANT: Sodoyer, Regis

; TITLE OF INVENTION: gp41 antigen
; FILE REFERENCE: 01-1692-A
; CURRENT APPLICATION NUMBER: US/10/263,103
; CURRENT FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 5312
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: plasmid pET-cer
US-10-263-103-35

Query Match 36.3%; Score 1199.2; DB 4; Length 5312;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGCGATAGAGCCCGGAAGAGA 60
DB 4622 CGGACACCATCGAATGGCGCAAAACCTTTCCGGGTATGCGATAGAGCCCGGAAGAGA 4563

QY 61 GTCGAATTCAGGGTGTGAATGTGAACACCAAGTAACCTTATACGATGTGCGAGAGTATGCCG 120
DB 4562 GTCGAATTCAGGGTGTGAATGTGAACACCAAGTAACCTTATACGATGTGCGAGAGTATGCCG 4503

QY 121 GTGTCCTTATACGACCGCTTTCCCGCGTGGTGAACACCGCCAGCCACGTTTCTGCGAAAA 180
DB 4502 GTGTCCTTATACGACCGCTTTCCCGCGTGGTGAACACCGCCAGCCACGTTTCTGCGAAAA 4443

QY 181 CGCGGGAAGAGTGAAGCGGGGATGGCGGAGCTGAATTACATTCCTCCAAACCGCTGGGCAC 240
DB 4442 CGCGGGAAGAGTGAAGCGGGGATGGCGGAGCTGAATTACATTCCTCCAAACCGCTGGGCAC 4383

QY 241 AACAACTCGCGGCAAAACAGTCTGTTGCTGATTGGGGTTTGCACCTCCAGTCTGGCCCTGC 300
DB 4382 AACAACTCGCGGCAAAACAGTCTGTTGCTGATTGGGGTTTGCACCTCCAGTCTGGCCCTGC 4323

QY 301 ACGCCCGCTCGCAAAATTTGCGCGCGATTAATCTCGCGCGATCAATCACTGGGTGCCAGCG 360
DB 4322 ACGCCCGCTCGCAAAATTTGCGCGCGATTAATCTCGCGCGATCAATCACTGGGTGCCAGCG 4263

QY 361 TGGTGGTTCGATGTAGAACGAAGCGGCGTGAAGCGCTGTAAAGCGGCGTGCACAATC 420
DB 4262 TGGTGGTTCGATGTAGAACGAAGCGGCGTGAAGCGCTGTAAAGCGGCGTGCACAATC 4203

QY 421 TTCTCGCGCAACGCTGAGTGGGCTGATCATTAATCTCCGCTGATGACACAGGATGCCA 480
DB 4202 TTCTCGCGCAACGCTGAGTGGGCTGATCATTAATCTCCGCTGATGACACAGGATGCCA 4143

QY 481 TTGCTGTGGAAGCTGCCTGCACATAATGTTTCGGCGGTTATTTCTGATGTCTTGACCCAGA 540
DB 4142 TTGCTGTGGAAGCTGCCTGCACATAATGTTTCGGCGGTTATTTCTGATGTCTTGACCCAGA 4083

QY 541 CACCCATCAACGATATTATTTCTCCATGAAGACGGTACGGCACTGGCGGTGGAGCATC 600
DB 4082 CACCCATCAACGATATTATTTCTCCATGAAGACGGTACGGCACTGGCGGTGGAGCATC 4023

QY 601 TGGTGCATTTGGTCCACGACCAATCGCGCTGTAGCGGGCCCATTAAGTTCTGCTCGG 660
DB 4022 TGGTGCATTTGGTCCACGACCAATCGCGCTGTAGCGGGCCCATTAAGTTCTGCTCGG 3963

QY 661 CCGCTCTGCGTCTGGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 3962 CCGCTCTGCGTCTGGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 3903

QY 721 CGGAACGGGAACGGCACTGGAGTGCATGTCCGGTTTCAACAAACCATGCAAAATGCTGA 780
DB 3902 CGGAACGGGAACGGCACTGGAGTGCATGTCCGGTTTCAACAAACCATGCAAAATGCTGA 3843

QY 781 ATGAGGGCATCTTCCCACTGCGATGCTGGTTCACACGATCAGATGGCGTGGCGGCA 840
DB 3842 ATGAGGGCATCTTCCCACTGCGATGCTGGTTCACACGATCAGATGGCGTGGCGGCA 3783

QY 841 TGCAGCCCATTTACCGAGTCCGGGCTCGCGGTGGTGGCGGATATCTCGGTAGTGGGATACG 900
DB 3782 TGCAGCCCATTTACCGAGTCCGGGCTCGCGGTGGTGGCGGATATCTCGGTAGTGGGATACG 3723

QY 901 ACGATACCGGAAGACAGCTCATGTTTATATCCCGCGGTTAAACCAACCATCAAAACAGGATTTTC 960
DB 3722 ACGATACCGGAAGACAGCTCATGTTTATATCCCGCGGTTAAACCAACCATCAAAACAGGATTTTC 3663

QY 961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTGCTCAACTCTCTCAGGGCCAGGCGGTGA 1020
DB 3662 GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTGCTCAACTCTCTCAGGGCCAGGCGGTGA 3603

QY 1021 AGGCAATCAGCTGTGTCCTGCTCACTGTGTAAGAAAGAAACCAACCTGCGGCCCAATA 1080
DB 3602 AGGCAATCAGCTGTGTCCTGCTCACTGTGTAAGAAAGAAACCAACCTGCGGCCCAATA 3543

QY 1081 CGCAAAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAATGAGTGGCAGCAGAGGTTT 1140
DB 3542 CGCAAAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAATGAGTGGCAGCAGAGGTTT 3483

QY 1141 CCCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
DB 3482 CCCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 3423

QY 1201 GCAC 1204
DB 3422 GCAC 3419

RESULT 12
US-08-929-967-1/c
; Sequence 1, Application US/089299967
; Patent No. 5891637
; GENERAL INFORMATION:
; APPLICANT: Rudpert, Siegfried J.W.
; TITLE OF INVENTION: Construction of Full-length cDNA Libraries
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/929,967
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1035R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5443 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-929-967-1

Query Match 36.3%; Score 1199.2; DB 2; Length 5443;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;


```
QY 1 CCGACACCATCGAATGTCGCAAAACCTTTTCGGGTATGGCATATAGCGCCCGGAAGAGA 60
Db 4810 CCGACACCATCGAATGTCGCAAAACCTTTTCGGGTATGGCATATAGCGCCCGGAAGAGA 4751
QY 61 GTCAATTCAGGGTGGTGAATGTCGAACCAAGTACGTTATACGATGTCGAGAGTATGCCG 120
Db 4750 GTCAATTCAGGGTGGTGAATGTCGAACCAAGTACGTTATACGATGTCGAGAGTATGCCG 4691
QY 121 GTGTCTCTTATCAGACCGTTTCCTCGGCGTGTGAACGAGCCGACGCTTTCTGCGAATA 180
Db 4690 GTGTCTCTTATCAGACCGTTTCCTCGGCGTGTGAACGAGCCGACGCTTTCTGCGAATA 4631
QY 181 CGCGGGAATAAGTGAAGCGCGCATGGCGAGTGAATTCACATTCCTCAACCGCTGGCAC 240
Db 4630 CGCGGGAATAAGTGAAGCGCGCATGGCGAGTGAATTCACATTCCTCAACCGCTGGCAC 4571
QY 241 AACAACTGGCGGCAACAGTGTCTCTGATTTGGGCTGGCGACCTCCAGTCTGGCCCTGC 300
Db 4570 AACAACTGGCGGCAACAGTGTCTCTGATTTGGGCTGGCGACCTCCAGTCTGGCCCTGC 4511
QY 301 AGCGCGCTGCGAAATTTGTGCGGCGAATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db 4510 AGCGCGCTGCGAAATTTGTGCGGCGAATTAATCTCGCGCGATCAACTGGGTGCCAGCG 4451
QY 361 TGGTGGTGCATGTTAGAACGAGCGCGTGAAGCTGTAAAGCTGTAAAGCGCGGTGCAATC 420
Db 4450 TGGTGGTGCATGTTAGAACGAGCGCGTGAAGCTGTAAAGCGCGGTGCAATC 4391
QY 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCAATTAATCTATCGGCTGATCACCAGGATGCA 480
Db 4390 TTCTCGCGCAACGCGTCAGTGGGCTGATCAATTAATCTATCGGCTGATCACCAGGATGCA 4331
QY 481 TTGCTGTGAAGCTGCTGCAATAATGTTTCGGCGTATTTCTTGATGTCTCTGACGAGA 540
Db 4330 TTGCTGTGAAGCTGCTGCAATAATGTTTCGGCGTATTTCTTGATGTCTCTGACGAGA 4271
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGGACTGGCGGTGAGCATC 600
Db 4270 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGGACTGGCGGTGAGCATC 4211
QY 601 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCTGTCTCGG 660
Db 4210 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCTGTCTCGG 4151
QY 661 CCGCTCTGCTGCTGCTGGCATTAATCTCACTCGGCAATCAAAATCAGCCGATAG 720
Db 4150 CCGCTCTGCTGCTGCTGGCATTAATCTCACTCGGCAATCAAAATCAGCCGATAG 4091
QY 721 CGAAACGGGAAGGCACTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 4090 CGAAACGGGAAGGCACTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 4031
QY 781 ATGAGGGCATGTTCCTCACTGCGATGCTGTTGTCGCAACGATCAGATGGCGTGGCGCAA 840
Db 4030 ATGAGGGCATGTTCCTCACTGCGATGCTGTTGTCGCAACGATCAGATGGCGTGGCGCAA 3971
QY 841 TGGCGCCATTAACGAGTTCGGGCTGGCGTGGTGGGATATCTCGTATGCGGATACG 900
Db 3970 TGGCGCCATTAACGAGTTCGGGCTGGCGTGGTGGGATATCTCGTATGCGGATACG 3911
QY 901 ACGATACGGAAGACAGTCTATGTTATATCCCGCGTTTAAACCCATCAAAACAGGATTTTC 960
Db 3910 ACGATACGGAAGACAGTCTATGTTATATCCCGCGTTTAAACCCATCAAAACAGGATTTTC 3851
QY 961 GCCTGCTGGGGCAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCAAGCGGTGA 1020
Db 3850 GCCTGCTGGGGCAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCAAGCGGTGA 3791
QY 1021 AGGCAATACAGTGTTCCTGCTCTCACTGGTGAAGAAAAAACACCCCTGGCGCCCAATA 1080
Db 3790 AGGCAATACAGTGTTCCTGCTCTCACTGGTGAAGAAAAAACACCCCTGGCGCCCAATA 3731
QY 1081 CGAAACCGGCTCTCCCGCGGCTTGGCCGATTAATATGACGTGGCACGACGAGTTT 1140
```

```
Db 3730 CGCAAAACCGCTCTCCCGCGCTTGGCCGATTCATTAATGACGTGCGACGACAGGTTT 3671
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTGAGTCTAGCTCAGTCATAG 1200
Db 3670 CCCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTGAGTCTAGCTCAGTCATAG 3611
QY 1201 GCAC 1204
Db 3610 GCAC 3607

RESULT 13
US-09-702-705-785/c
; Sequence 785, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-785

Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGTCGCAAAACCTTTTCGGGTATGGCATATAGCGCCCGGAAGAGA 60
Db 4677 CCGACACCATCGAATGTCGCAAAACCTTTTCGGGTATGGCATATAGCGCCCGGAAGAGA 4618
QY 61 GTCAATTCAGGGTGGTGAATGTGAACCAAGTAAACGTTATACGATGTCGAGAGTATGCCG 120
Db 4617 GTCAATTCAGGGTGGTGAATGTGAACCAAGTAAACGTTATACGATGTCGAGAGTATGCCG 4558
QY 121 GTGTCTCTTATCAGACCGTTTCCTCGGCGTGTGAACGAGCCGACGCTTTCTGCGAATA 180
Db 4557 GTGTCTCTTATCAGACCGTTTCCTCGGCGTGTGAACGAGCCGACGCTTTCTGCGAATA 4498
QY 181 CGCGGAAAAAGTGAAGCGCGCATGGCGGAGTGAATTCACATTCCTCAACCGCGTGGCAC 240
Db 4497 CGCGGAAAAAGTGAAGCGCGCATGGCGGAGTGAATTCACATTCCTCAACCGCGTGGCAC 4438
QY 241 AACAACTGGCGGCAACAGTGTGCTGATTTGGGCTGGCGACCTCCAGTCTGGCCCTGC 300
Db 4437 AACAACTGGCGGCAACAGTGTGCTGATTTGGGCTGGCGACCTCCAGTCTGGCCCTGC 4378
QY 301 AGCGCGCTGCGAAATTTGTGCGGCGGATTTAAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db 4377 AGCGCGCTGCGAAATTTGTGCGGCGGATTTAAATCTCGCGCGATCAACTGGGTGCCAGCG 4318
QY 361 TGGTGGTGTGATGTGTAGAAAGAGCGGCTGCAAGCTGTAAAGCGCGGTGCACAATC 420
Db 4317 TGGTGGTGTGATGTGTAGAAAGAGCGGCTGCAAGCTGTAAAGCGCGGTGCACAATC 4258
QY 421 TTCTCGCGCAACGCGTCAAGTGGGCTGATCAATTAATCTCCGCTCGGTGATGCAACGAGTTGCA 480
```

Db 4257 TTCTCGCCAAAGCGTCTAGTGGCTGATCATTAACATATCCGCTGGATGACCAAGATGCCA 4198
Qy 481 TTGCTGTGGAGCTGCCTGCACTAAATGTTCCGGCGTTATTTCTTGGATCTCTGACCAAGA 540
Db 4197 TTGCTGTGGAGCTGCCTGCACTAAATGTTCCGGCGTTATTTCTTGGATCTCTGACCAAGA 4138
Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGCTACCGACTGGCGTGGAGCATC 600
Db 4137 CACCCATCAACAGTATTATTTCTCCCATGAAGACGCTACCGACTGGCGTGGAGCATC 4078
Qy 601 TGGTCCGATTGGTTCACAGCAAAATCGCGCTGTTAGCGGGCCCAATAAATGTTCTGTTCTCGG 660
Db 4077 TGGTCCGATTGGTTCACAGCAAAATCGCGCTGTTAGCGGGCCCAATAAATGTTCTGTTCTCGG 4018
Qy 661 CGCGTCTCGCTGCTGGCTGGCGATAAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4017 CGCGTCTCGCTGCTGGCTGGCGATAAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 3958
Qy 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTCACAAACCAATGCNAATGCTGA 780
Db 3957 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTCACAAACCAATGCNAATGCTGA 3898
Qy 781 ATGAGGGCATGTTCCCACTGCGATGCTGTTGCCAAGATCAGATGCGCTGGGCGCAA 840
Db 3897 ATGAGGGCATGTTCCCACTGCGATGCTGTTGCCAAGATCAGATGCGCTGGGCGCAA 3838
Qy 841 TCGCGCCCATACCAAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db 3837 TCGCGCCCATACCAAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 3778
Qy 901 ACGATACGGAAGACAGCTCATGTTATATCCCGCGGTTTAAACCAATCAAAACGAGATTTC 960
Db 3777 ACGATACGGAAGACAGCTCATGTTATATCCCGCGGTTTAAACCAATCAAAACGAGATTTC 3718
Qy 961 GCCTCTGGGGCAAAACAGCTGGACCGCTTCTGCTCAACTCTCTCAGGCGCAGGCGGTGA 1020
Db 3717 GCCTCTGGGGCAAAACAGCTGGACCGCTTCTGCTCAACTCTCTCAGGCGCAGGCGGTGA 3658
Qy 1021 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAGAAACCAACCTGGCGGCCAATA 1080
Db 3657 AGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAGAAACCAACCTGGCGGCCAATA 3598
Qy 1081 CGCAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGAGCTGGCAGCAGAGGTTT 1140
Db 3597 CGCAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGAGCTGGCAGCAGAGGTTT 3538
Qy 1141 CCCGACTGGAAGCGGCGAGTGCAGCGCAACCAATTAATGAGCTGCTCACTCATTTAG 1200
Db 3537 CCCGACTGGAAGCGGCGAGTGCAGCGCAACCAATTAATGAGCTGCTCACTCATTTAG 3478
Qy 1201 GCAC 1204
Db 3477 GCAC 3474

RESULT 14

US-09-736-457-785/c
; Sequence 785, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Jingtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Pan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15

; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-785

Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGFTGCAAAACCTTTTCGGGTATGCGCATGATGAGCCCGGAAGAGA 60
Db 4677 CCGACACCATCGAATGFTGCAAAACCTTTTCGGGTATGCGCATGATGAGCCCGGAAGAGA 4618
Qy 61 GTCAATTCAGGGTGGTGAATGTAAGCAAGTAAAGTTATACGATGTCGACAGTATGCG 120
Db 4617 GTCAATTCAGGGTGGTGAATGTAAGCAAGTAAAGTTATACGATGTCGACAGTATGCG 4558
Qy 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACAGGCCAGCCACGTTTCTGCGAAAA 180
Db 4557 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACAGGCCAGCCACGTTTCTGCGAAAA 4498
Qy 181 CGCGGAAAAAGTGGAGCGCGATGCGGAGCTGAATTAATTCCCAACCGCGTGGCAC 240
Db 4497 CGCGGAAAAAGTGGAGCGCGATGCGGAGCTGAATTAATTCCCAACCGCGTGGCAC 4438
Qy 241 AACAACTGGCGGGCAAAACAGTCTGTTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 4437 AACAACTGGCGGGCAAAACAGTCTGTTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGC 4378
Qy 301 ACGCGCCCTCGCAAAATGTCGCGCGCATTAATAATCTCGCGCGCATCAACTGGGTGCCAGCG 360
Db 4377 ACGCGCCCTCGCAAAATGTCGCGCGCATTAATAATCTCGCGCGCATCAACTGGGTGCCAGCG 4318
Qy 361 TGGTGGTTCGATGGTGAAGCGAGCGCGCTGGAAGCTGTAAGCGCGCGGTGCAAAATC 420
Db 4317 TGGTGGTTCGATGGTGAAGCGAGCGCGCTGGAAGCTGTAAGCGCGCGGTGCAAAATC 4258
Qy 421 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATATCCGCTGGATGACCAAGGATGCCA 480
Db 4257 TTCTCGCGCAACGGCTCAGTGGGCTGATCAATTAATATCCGCTGGATGACCAAGGATGCCA 4198
Qy 481 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTTGACCCAGA 540
Db 4197 TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTTCTTGACCCAGA 4138
Qy 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGCTACGCGACTGGGCGTGGAGCATC 600
Db 4137 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGCTACGCGACTGGGCGTGGAGCATC 4078
Qy 601 TGGTCCGATTGGTTCACAGCAAAATCGCGCTGTTAGCGGGCCCAATAAATGTTCTGTTCTCGG 660
Db 4077 TGGTCCGATTGGTTCACAGCAAAATCGCGCTGTTAGCGGGCCCAATAAATGTTCTGTTCTCGG 4018
Qy 661 CGCGTCTCGCTGCTGGCTGGCGATAAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4017 CGCGTCTCGCTGCTGGCTGGCGATAAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 3958
Qy 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTCACAAACCAATGCNAATGCTGA 780
Db 3957 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTCACAAACCAATGCNAATGCTGA 3898
Qy 781 ATGAGGGCATGTTCCCACTGCGATGCTGTTGCCAAGATCAGATGCGCTGGGCGCAA 840
Db 3897 ATGAGGGCATGTTCCCACTGCGATGCTGTTGCCAAGATCAGATGCGCTGGGCGCAA 3838
Qy 841 TCGCGCCCATACCAAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db 3837 TCGCGCCCATACCAAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG 3778

```
QY 901 ACATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCACTCAACAGGATTTTC 960
Db 3777 ACATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCACTCAACAGGATTTTC 3718
QY 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGGGTGA 1020
Db 3717 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGGGTGA 3658
QY 1021 AGGGCAATCAGCTGTGCGCGCTCTACTGTTGAAAGAAAACCACTTGGCGCCCAATA 1080
Db 3657 AGGGCAATCAGCTGTGCGCGCTCTACTGTTGAAAGAAAACCACTTGGCGCCCAATA 3598
QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTATGAGTGGCAGCAGAGTTT 1140
Db 3597 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTATGAGTGGCAGCAGAGTTT 3538
QY 1141 CCGGACTGAAAGCGGGGAGTGGCGCAACGCAATTAATGAGTGGTACTCACTATTAG 1200
Db 3537 CCGGACTGAAAGCGGGGAGTGGCGCAACGCAATTAATGAGTGGTACTCACTATTAG 3478
QY 1201 GCAC 1204
Db 3477 GCAC 3474

RESULT 15
US-09-614-124B-785/c
; Sequence 785, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-785

Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGTGCAAAACCTTTCGGGTATGGCATGATAGCGCCCGAAGAGA 60
Db 4677 CCGACACCATCGAATGTGCGCAAAACCTTTCGGGTATGGCATGATAGCGCCCGAAGAGA 4618
QY 61 GTCAATT CAGGGTGGTGAATGTGAACACAGTAACTGTTATACGATGTCGCGAGATGTCGG 120
Db 4617 GTCAATTCAGGGTGGTGAATGTGAACACAGTAACTGTTATACGATGTCGCGAGATGTCGG 4558
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCTGTGTAACACGAGCCAGCCAGTTTCTCGCAAAA 180
Db 4557 GTGTCTCTTATCAGACCGTTTCCCGCTGTGTAACACGAGCCAGCCAGTTTCTCGCAAAA 4498
QY 181 CCGGGGAAAAGTGGAGCGCGATGCGGAGCTGAATTACATTCCTCCAAACCGCGTGGCAC 240
Db 4497 CCGGGGAAAAGTGGAGCGCGATGCGGAGCTGAATTACATTCCTCCAAACCGCGTGGCAC 4438
QY 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGGCCACTCCAGCTCTGGCCCTGC 300
```

```
Db 4437 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGCACCTCCAGTCGTGGCCCTGC 4378
QY 301 ACGGCGCGTCCCAAAATTTGTTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
Db 4377 ACGGCGCGTCCCAAAATTTGTTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 4318
QY 361 TGGTGGTGTGATCGGTGAGAACGAGCGCGTTCGAGGCTGTAAAGCGCGGTGCACAATC 420
Db 4317 TGGTGGTGTGATCGGTGAGAACGAGCGCGTTCGAGGCTGTAAAGCGCGGTGCACAATC 4258
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTA 480
Db 4257 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTA 4198
QY 481 TTCTGTGGAAGCTGCTGCACCTAATGTTCCGGCGTTAATTTCTTGTGATGTTCTGTGACCGA 540
Db 4197 TTCTGTGGAAGCTGCTGCACCTAATGTTCCGGCGTTAATTTCTTGTGATGTTCTGTGACCGA 4138
QY 541 CACCCATCAACAGTATTAATTTCTCCATGAAGCGGTA CGCGACTGGGCGTGGAGCATC 600
Db 4137 CACCCATCAACAGTATTAATTTCTCCATGAAGCGGTA CGCGACTGGGCGTGGAGCATC 4078
QY 601 TGGTGGCATTTGGGTGTCACGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCGG 660
Db 4077 TGGTGGCATTTGGGTGTCACGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCGG 4018
QY 661 CGCGTCTGCGTCTGGGTGGTGGCATAAATPATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4017 CGCGTCTGCGTCTGGGTGGTGGCATAAATPATCTCACTCGCAATCAAAATTCAGCCGATAG 3958
QY 721 CGGAACGGGAAGCGGCTGAGTGGCATGTCGGTTCGCAAAACGATAGTGGCGCTGGCGCAA 840
Db 3957 CGGAACGGGAAGCGGCTGAGTGGCATGTCGGTTCGCAAAACGATAGTGGCGCTGGCGCAA 3898
QY 781 ATGAGGCGCATCGTTCCCACTCGCATGTCGGTTCGCAAAACGATAGTGGCGCTGGCGCAA 840
Db 3897 ATGAGGCGCATCGTTCCCACTCGCATGTCGGTTCGCAAAACGATAGTGGCGCTGGCGCAA 3838
QY 841 TGGCGGCATTCAGGATCGCGCTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 900
Db 3837 TGGCGGCATTCAGGATCGCGCTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 3778
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTACCAACCATCAACAGGATTTTC 960
Db 3777 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTACCAACCATCAACAGGATTTTC 3718
QY 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGCGTGA 1020
Db 3717 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGCGTGA 3658
QY 1021 AGGGCAATCAGCTGTGCGCGCTCTCACTGGTGAAGAAAACCACTTGGCGCCCAATA 1080
Db 3657 AGGGCAATCAGCTGTGCGCGCTCTCACTGGTGAAGAAAACCACTTGGCGCCCAATA 3598
QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTATGAGTGGCAGCAGAGTTT 1140
Db 3597 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTATGAGTGGCAGCAGAGTTT 3538
QY 1141 CCGGACTGAAAGCGGGGAGTGGCGCAACGCAATTAATGAGTGGTACTCACTATTAG 1200
Db 3537 CCGGACTGAAAGCGGGGAGTGGCGCAACGCAATTAATGAGTGGTACTCACTATTAG 3478
QY 1201 GCAC 1204
Db 3477 GCAC 3474
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 1309.2 Seconds

(without alignments)
16314.373 Million cell updates/sec

Title: US-09-765-555B-18

Perfect score: 3300

Sequence: 1 ccgacacatcgatgtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3300	100.0	3300	10	US-09-765-555-18	Sequence 18, Appl
2	3232.8	98.0	3300	10	US-09-765-555-17	Sequence 17, Appl
3	3226.4	97.8	3300	10	US-09-765-555-14	Sequence 15, Appl
4	3226.4	97.8	3300	10	US-09-765-555-15	Sequence 14, Appl
5	3216.8	97.5	3300	10	US-09-765-555-16	Sequence 16, Appl
6	2705.8	82.0	7373	18	US-10-257-384A-1	Sequence 1, Appl
7	2705.8	82.0	8101	17	US-10-149-472-5	Sequence 5, Appl

8	2701	81.8	6648	19	US-10-343-859-8	Sequence 8, Appl
9	2698	81.8	9191	19	US-10-343-859-9	Sequence 9, Appl
10	2690.4	81.5	7112	18	US-10-263-153-40	Sequence 40, Appl
11	2690.4	81.5	7259	18	US-10-263-153-30	Sequence 30, Appl
12	2690.4	81.5	7322	18	US-10-263-153-35	Sequence 35, Appl
13	2690.4	81.5	7352	18	US-10-263-153-25	Sequence 25, Appl
14	2690.4	81.5	7370	18	US-10-263-153-20	Sequence 20, Appl
15	2690.4	81.5	7370	18	US-10-263-153-61	Sequence 61, Appl
16	2690.4	81.5	7370	18	US-10-263-153-66	Sequence 66, Appl
17	2690.4	81.5	7370	18	US-10-263-153-71	Sequence 71, Appl
18	2690.4	81.5	7403	18	US-10-263-153-15	Sequence 15, Appl
19	2690.4	81.5	7442	18	US-10-263-153-10	Sequence 10, Appl
20	2690.4	81.5	7478	18	US-10-263-153-3	Sequence 3, Appl
21	2636.8	79.9	6806	10	US-09-908-943A-194	Sequence 194, App
22	2636.8	79.9	6806	20	US-10-801-487-194	Sequence 194, App
23	2636.8	79.9	6806	20	US-10-801-938-194	Sequence 194, App
24	2636.8	79.9	6806	20	US-10-801-509-194	Sequence 194, App
25	2636.8	79.9	6806	20	US-10-801-486-194	Sequence 194, App
26	2636.8	79.9	6806	21	US-10-801-493-194	Sequence 194, App
27	2605.4	79.0	7553	18	US-10-263-153-7	Sequence 7, Appl
28	1609.4	48.8	5558	16	US-10-241-596-137	Sequence 137, App
29	1474	44.7	4700	17	US-10-313-963A-54	Sequence 54, Appl
30	1462.4	44.3	4921	21	US-10-495-491-2	Sequence 2, Appl
31	1462.4	44.3	4935	21	US-10-495-491-1	Sequence 1, Appl
32	1462.4	44.3	4945	21	US-10-495-491-3	Sequence 3, Appl
33	1462.4	44.3	4951	21	US-10-495-491-4	Sequence 4, Appl
34	1417.6	43.0	1511	10	US-09-987-763-30	Sequence 30, Appl
C 35	1417.6	43.0	5926	13	US-10-024-809-3	Sequence 3, Appl
C 36	1238.4	37.5	5024	17	US-10-313-963A-52	Sequence 52, Appl
C 37	1238.4	37.5	5448	21	US-10-671-995A-2	Sequence 2, Appl
C 38	1238.4	37.5	5454	21	US-10-671-995A-1	Sequence 1, Appl
39	1237.2	37.5	1922	10	US-09-995-998A-32	Sequence 32, Appl
40	1237.2	37.5	1922	17	US-10-420-034A-32	Sequence 32, Appl
41	1237.2	37.5	6270	21	US-10-855-897-2	Sequence 2, Appl
42	1204	36.5	4557	17	US-10-457-372-5	GENERAL INFORMA
43	1204	36.5	6664	18	US-10-270-176-1	Sequence 1, Appl
44	1204	36.5	6668	18	US-10-270-176-2	Sequence 2, Appl
45	1204	36.5	6742	18	US-10-270-176-19	Sequence 19, Appl

ALIGNMENTS

RESULT 1
US-09-765-555-18
; Sequence 18, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Parial sequence of pMal- λ p3 and zinc finger
; OTHER INFORMATION: protein ZFPap3
US-09-765-555-18

Query Match 100.0%; Score 3300; DB 10; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTCAAAACCTTTGCGGGTATGGCATGTAGCGCCCGGAAGA 60
Db 1 CCGACACCATCGAATGGTCAAAACCTTTGCGGGTATGGCATGTAGCGCCCGGAAGA 60
QY 61 GTCAAATTCAGGTCGTGAATGTGAACACGATGAACGTTATACGATGCGCAGAGTATGCCG 120
Db 61 GTCAAATTCAGGTCGTGAATGTGAACACGATGAACGTTATACGATGCGCAGAGTATGCCG 120
QY 121 GTGTCTCTTTACAGACCGGTTTCCGCGTGTGTGAACCGAGCCGACGACGTTTCTGCGAAAA 180
Db 121 GTGTCTCTTTACAGACCGGTTTCCGCGTGTGTGAACCGAGCCGACGACGTTTCTGCGAAAA 180
QY 181 CCGCGGAAAAAGTGGAAAGCGGCGATGGCGAGCTGAAATTCACATCCCAACCGCGTGGCAC 240
Db 181 CCGCGGAAAAAGTGGAAAGCGGCGATGGCGAGCTGAAATTCACATCCCAACCGCGTGGCAC 240
QY 241 AACCACTGGCGGCAACAGTGGTGTGATTTGGGGTGTGCCACCTCCAGTCTGGGCCCTGC 300
Db 241 AACCACTGGCGGCAACAGTGGTGTGATTTGGGGTGTGCCACCTCCAGTCTGGGCCCTGC 300
QY 301 ACGCGCGTCCGCAAAATTTGTCCGGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGTCCGCAAAATTTGTCCGGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTGTGATGGTAGAACGAGCGGCTCGAAGCCGTGTAAAGCGGCGGTGCAACAATC 420
Db 361 TGGTGGTGTGATGGTAGAACGAGCGGCTCGAAGCCGTGTAAAGCGGCGGTGCAACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTCCGCTGATCACCAGNATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTCCGCTGATCACCAGNATGCCA 480
QY 481 TTGCTGTGGAGCTGCCTGCATTAATGTTCCGGCGTATTTCTTGATGTCTCTGACCAGA 540
Db 481 TTGCTGTGGAGCTGCCTGCATTAATGTTCCGGCGTATTTCTTGATGTCTCTGACCAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGATCGGGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGATCGGGGTGGAGCATC 600
QY 601 TGGTCCGATTTGGGTCAACAGAAATCGCGCTGTTAGCGGGGCCATTAAGTTCTGTCCTCG 660
Db 601 TGGTCCGATTTGGGTCAACAGAAATCGCGCTGTTAGCGGGGCCATTAAGTTCTGTCCTCG 660
QY 661 CCGCTCTGGCTGTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CCGCTCTGGCTGTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGACCTGGAGTGCATGTCGGTTCACAAACCAATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGACCTGGAGTGCATGTCGGTTCACAAACCAATGCAAAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTGTGCCAAGATCAGATGCGCTGGCGGCA 840
Db 781 ATGAGGGCATGTTCCCACTGCGATGCTGGTGTGCCAAGATCAGATGCGCTGGCGGCA 840
QY 841 TGGCGGCCATTCAGAGTCCGGCTGCGCTGTGGTGGCGATATCTCGTAGTGGGATACG 900
Db 841 TGGCGGCCATTCAGAGTCCGGCTGCGCTGTGGTGGCGATATCTCGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTAAACCAACCATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTAAACCAACCATCAACAGGATTTTC 960
QY 961 GCCTGTGGGGCAACACGAGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGTGGGGCAACACGAGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGGCCGTCTCATCTGGTGAAGAAAGAAACCAACCTCGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCGTCTCATCTGGTGAAGAAAGAAACCAACCTCGCGCCCAATA 1080
QY 1081 CGCAAAACCGCCTCTCCCGCGCGCTTGGCCGATTCATTAATGCAGCTGCGCAGCATTTT 1140

Db 1081 CGCAAAACCGCCTCTCCCGCGCGTTGGCCGATTCATTAATGCAGCTGCGCAGCAGGTTT 1140
QY 1141 CCGAGTGGAAAGCGGCGAGTGAAGCGCAACGCAATTAATGTAGTGTAGCTCACTCATTTAG 1200
Db 1141 CCGAGTGGAAAGCGGCGAGTGAAGCGCAACGCAATTAATGTAGTGTAGCTCACTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTGGACAGCTTATCATGCACTGCAAGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGGACAGCTTATCATGCACTGCAAGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACGCCATCCGAAAGCTGTGTATGGCTGTGCAGCTGTGAATCACTGCATAAATTCG 1320
Db 1261 TCAGGACGCCATCCGAAAGCTGTGTATGGCTGTGCAGCTGTGAATCACTGCATAAATTCG 1320
QY 1321 TGTGCTCAAGGCCACATCCCGTTCTCGATAAATGTTTTTTCGCCCGACATCATAAACGGTT 1380
Db 1321 TGTGCTCAAGGCCACATCCCGTTCTCGATAAATGTTTTTTCGCCCGACATCATAAACGGTT 1380
QY 1381 CTGGCAATAATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA 1440
Db 1381 CTGGCAATAATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA 1440
QY 1441 ATTGTGAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTATAGTGTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTATAGTGTTCACGA 1500
QY 1501 GCATTCACAAACAGGACCATAGATTAATGAAACTGGAAGAGTAACTGGTAAATCTCG 1560
Db 1501 GCATTCACAAACAGGACCATAGATTAATGAAACTGGAAGAGTAACTGGTAAATCTCG 1560
QY 1561 ATTAACGGCGATAAAGGCTAATACGGTCTCGCTGAACTGCGTAAAGAAATTCGAGAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTAATACGGTCTCGCTGAACTGCGTAAAGAAATTCGAGAAGAT 1620
QY 1621 ACCGGAATTAAGTACCGTTGAGCATCCGGAATAACTGGAAGAGAAATTCGCCACAGTT 1680
Db 1621 ACCGGAATTAAGTACCGTTGAGCATCCGGAATAACTGGAAGAGAAATTCGCCACAGTT 1680
QY 1681 GCGGCACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACAGCCGTTTGGTGGCTAC 1740
Db 1681 GCGGCACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACAGCCGTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACGGCAAGCTGATTCCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACGGCAAGCTGATTCCTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAACCTCGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAACCTCGGAA 1920
QY 1921 GAGATCCGCGCTGATGATAAGAACTGAAAGGAGAAAGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGCGCTGATGATAAGAACTGAAAGGAGAAAGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGGTACTTACCTTGGCGCTGATTCGCTGACGGGGTTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGGTACTTACCTTGGCGCTGATTCGCTGACGGGGTTTATGCGTTCAAG 2040
QY 2041 TATGAAACGGCAAGTACGACATTAAGAGCTGGGCGTGGATAACCGCTGGCGGAAGCG 2100
Db 2041 TATGAAACGGCAAGTACGACATTAAGAGCTGGGCGTGGATAACCGCTGGCGGAAGCG 2100
QY 2101 GGTCTGACCTTCTGCTGCTGACCTGATTAATAACCAACATGAATGCACACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGCTGACCTGATTAATAACCAACATGAATGCACACCGGATTAC 2160
QY 2161 TCCATCGCAAGAGCTGCCCTTTTAAAGGCGAAACAGCGATGACCAATCAACCGCCGCTGG 2220

Db 2161 TCCATCGCAGAGCTGCCCTTTAATAAAGCGGAAACAGCGATGATCACCATCAACGGCCCGGTGG 2220
Qy 2221 GCATGGTCAACATCGACACACGAGAAAGTGAATTAATGTTGTAACCGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACACGAGAAAGTGAATTAATGTTGTAACCGTACTGCGGACCTTC 2280
Qy 2281 AAGGGTCAACCATCAAAACCGTTCTGTCGCGTCTGAGCGAGGTATTAAGCGCGCCAGT 2340
Db 2281 AAGGGTCAACCATCAAAACCGTTCTGTCGCGTCTGAGCGAGGTATTAAGCGCGCCAGT 2340
Qy 2341 CCGAAACAAAGAGCTCGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGCTGTG 2400
Db 2341 CCGAAACAAAGAGCTCGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGCTGTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGTCCGTAGCGCTGAAGTCTTACGAGAAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGTCCGTAGCGCTGAAGTCTTACGAGAAAG 2460
Qy 2461 TTGGCGAAGATCCACGATATTGCCGCCACCATGGAAGACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGATATTGCCGCCACCATGGAAGACGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAAACATCCCGCAGATGTCGCTTCTTCTGGTATGCGGTGCTGCTGCTGATCAACGCC 2580
Db 2521 CCGAAACATCCCGCAGATGTCGCTTCTTCTGGTATGCGGTGCTGCTGCTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Qy 2641 AACACACACAAATAACAAATTAACAAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACACACACAAATAACAAATTAACAAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Qy 2701 GGATCCTCTTCTCTGTGGCCAGCGGCCCTCGAGCCCGGGAGAACCTTATGCTGTG 2760
Db 2701 GGATCCTCTTCTCTGTGGCCAGCGGCCCTCGAGCCCGGGAGAACCTTATGCTGTG 2760
Qy 2761 CCGGAATGTGTAAGTCTTTCAGCCAGAGCAGCTCCCTGGTGGCCACACAGCGTACCCAC 2820
Db 2761 CCGGAATGTGTAAGTCTTTCAGCCAGAGCAGCTCCCTGGTGGCCACACAGCGTACCCAC 2820
Qy 2821 ACGGGTGAACACCGTATAAATGCGCAGAGTGCAGAAATCTTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGGTGAACACCGTATAAATGCGCAGAGTGCAGAAATCTTTTAGCCAGTCCAGCAAC 2880
Qy 2881 CTGGTGGCCATCAACGCACTCATACTGGCGAGAGGCCATACAAATGTCCAGAAATGTGGC 2940
Db 2881 CTGGTGGCCATCAACGCACTCATACTGGCGAGAGGCCATACAAATGTCCAGAAATGTGGC 2940
Qy 2941 AAGTCTTTTCAGCCAGTCCAGCAACCTGTTGCGCCACCAACGTACTCACACCGGGGAGAG 3000
Db 2941 AAGTCTTTTCAGCCAGTCCAGCAACCTGTTGCGCCACCAACGTACTCACACCGGGGAGAG 3000
Qy 3001 CCCTATGCTTGTTCGGAAATGTGGTAAGTCTTTCAGCAACAGTGGCTCTCTTGTGTTAGACAC 3060
Db 3001 CCCTATGCTTGTTCGGAAATGTGGTAAGTCTTTCAGCAACAGTGGCTCTCTTGTGTTAGACAC 3060
Qy 3061 CAGCGTACCCACACCGGTGAAAACCGTATAAATGCCAGAGTGGGCAAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACCGGTGAAAACCGTATAAATGCCAGAGTGGGCAAAATCTTTTAGC 3120
Qy 3121 CAGCGCGCCACTTGGAAACCGCATCAACGCACTCATACTGCGGAGAGGCCATACAAATGT 3180
Db 3121 CAGCGCGCCACTTGGAAACCGCATCAACGCACTCATACTGCGGAGAGGCCATACAAATGT 3180
Qy 3181 CCAGAAATGTGCAAGTCTTTTCTCAACTTCAGGCAACTTGGTCCGTCACCAACGTAATCTAC 3240
Db 3181 CCAGAAATGTGCAAGTCTTTTCTCAACTTCAGGCAACTTGGTCCGTCACCAACGTAATCTAC 3240
Qy 3241 ACCGGTAAAAAATAGTGGCCAGCGCCGCGAGTACCCGTCAGAGTTCGGGACTACGCT 3300
Db 3241 ACCGGTAAAAAATAGTGGCCAGCGCCGCGAGTACCCGTCAGAGTTCGGGACTACGCT 3300

RESULT 2
US-09-765-555-17
; Sequence 17, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE REFERENCE: 27801-20014-40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m4 and zinc finger
; OTHER INFORMATION: protein ZFPm4
US-09-765-555-17

Query Match 98.0%; Score 3232.8; DB 10; Length 3300;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 3258; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGTGCAAAACCTTTTCGCGGTATGGCATGTAGGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGTGCAAAACCTTTTCGCGGTATGGCATGTAGGCGCCCGGAAGAGA 60
Qy 61 CTCGAATTCAGGGTGGTGAATGTGAACCAAGTAAAGTTATACGATGTCCAGAGTATGCCG 120
Db 61 CTCGAATTCAGGGTGGTGAATGTGAACCAAGTAAAGTTATACGATGTCCAGAGTATGCCG 120
Qy 121 GTGTCCTTTATCAGACCCGTTTCCCGCGTGTGAACACGAGCCAGCCAGCTTTCTCGGAAAA 180
Db 121 GTGTCCTTTATCAGACCCGTTTCCCGCGTGTGAACACGAGCCAGCCAGCTTTCTCGGAAAA 180
Qy 181 CCGCGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTAATATCCCAACCGCGTGGCAC 240
Db 181 CCGCGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTAATATCCCAACCGCGTGGCAC 240
Qy 241 AACAACTCGCGGGCAACAGTCGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTCGCGGGCAACAGTCGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGGCCGTTCGCAAAATTCGCGCGGATTAATCTCGCGCCGATCAACTGGGTGGCAGCG 360
Db 301 ACGGCCGTTCGCAAAATTCGCGCGGATTAATCTCGCGCCGATCAACTGGGTGGCAGCG 360
Qy 361 TGGTGGTTCGATGGTAGAACGAAAGCGCGTGGAAAGCCCTGTAAAGCGCGGTGCACAAATC 420
Db 361 TGGTGGTTCGATGGTAGAACGAAAGCGCGTGGAAAGCCCTGTAAAGCGCGGTGCACAAATC 420
Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATTAATATCCGCTGATGACCAAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATTAATATCCGCTGATGACCAAGGATGCCA 480
Qy 481 TTGCTGTGGAAGCTCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTCCCTGCACTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCCAGA 540
Qy 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGCTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGCTACGCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTCCGATTTGGGTCCACAGCAAAATCGCGCTGTTAGCGGCGCCCATTAAGTTCTGTCTCGG 660

Db 601 TGGTCGATTTGGGTCACAGCAAAATCGCGCTGTAGCGGCCCAITTAAGTTCTGTCTCGG 660
Qy 661 CGCGTCTGGCTCTGGCTGGCTGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGGCTCTGGCTGGCTGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGCGCACTGGAGTGCATGTCCGGTTTTCAACAAACCAATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGCGCACTGGAGTGCATGTCCGGTTTTCAACAAACCAATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATCTTCCCACTGCGATCTGGTTGCCAAACGATCAGATGGCGCTGGCGCAA 840
Db 781 ATGAGGGCATCTTCCCACTGCGATCTGGTTGCCAAACGATCAGATGGCGCTGGCGCAA 840
Qy 841 TGGCGGCCATTACCGAGTCCGGGCTCGCGGTTGGTGGGATATCTCGTGGGATACG 900
Db 841 TGGCGGCCATTACCGAGTCCGGGCTCGCGGTTGGTGGGATATCTCGTGGGATACG 900
Qy 901 AGGATACCGAAGACAGCTCATGTTATATCCGCCGTTAACCAACCAAGGATTTTC 960
Db 901 AGGATACCGAAGACAGCTCATGTTATATCCGCCGTTAACCAACCAAGGATTTTC 960
Qy 961 GCCTGCTGGGGCAACCAAGCGTGGACCGCTTCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1020
Db 961 GCCTGCTGGGGCAACCAAGCGTGGACCGCTTCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1020
Qy 1021 AGGGCAATCAGCTGTTGCCGCTTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCGCTTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGTGGCACGACAGTTT 1140
Db 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGTGGCACGACAGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTAACTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTAACTCACTCATTTAG 1200
Qy 1201 GCACAAATCTCATGTTTGAACAGCTTATCATCGACTGCAACGCTGACCAATCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGAACAGCTTATCATCGACTGCAACGCTGACCAATCTTCTGGCG 1260
Qy 1261 TCAGGACGCAATCGGAAGCTGTTGATGCTGTCGAGTGTGAATCACTGCAATTCG 1320
Db 1261 TCAGGACGCAATCGGAAGCTGTTGATGCTGTCGAGTGTGAATCACTGCAATTCG 1320
Qy 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATTAATGTTTTTTCGCGGACATCATACCGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATTAATGTTTTTTCGCGGACATCATACCGGTT 1380
Qy 1381 CTGGCAATATCTGAAATGAGCTGTTGA CAATTAATCACTCGGCTCGTATAATGTTGGA 1440
Db 1381 CTGGCAATATCTGAAATGAGCTGTTGA CAATTAATCACTCGGCTCGTATAATGTTGGA 1440
Qy 1441 ATTGTGAGCGGATAACCAATTTTCAACAGGAACAGCGAGTAACTGGTAACTCTGG 1500
Db 1441 ATTGTGAGCGGATAACCAATTTTCAACAGGAACAGCGAGTAACTGGTAACTCTGG 1500
Qy 1501 GCACTTCAACCAAGGACCATAGATTTATGAAGTGAAGGTAACCTGGTAACTCTGG 1560
Db 1501 GCACTTCAACCAAGGACCATAGATTTATGAAGTGAAGGTAACCTGGTAACTCTGG 1560
Qy 1561 ATTAACCGGATTAAGGCTATTAACGGTCTCGCTGAAGTGGTAACTGGTAACTCTGG 1620
Db 1561 ATTAACCGGATTAAGGCTATTAACGGTCTCGCTGAAGTGGTAACTGGTAACTCTGG 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGAAGAGAAATTCGAGAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGAAGAGAAATTCGAGAGGTT 1680
Qy 1681 CGGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACGACCGCTTTGGTGGCTAC 1740

Db 1681 GCGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGCGCTGTTGGCTGGAATCACCCCGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTTGGCTGGAATCACCCCGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACCTCGGATGCGGTACGTTCAACGGCAAGCTGATTGCTTACCGATCGCTGTT 1860
Db 1801 CCGTTTACCTCGGATGCGGTACGTTCAACGGCAAGCTGATTGCTTACCGATCGCTGTT 1860
Qy 1861 GAACGCTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCGCCCAAAACCTCGGAA 1920
Db 1861 GAACGCTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCGCCCAAAACCTCGGAA 1920
Qy 1921 GAGATCCGGCGCTGGATTAAGAACTGAAAGCGCAAGGTAAGAGCGCGCTGATTTCAAC 1980
Db 1921 GAGATCCGGCGCTGGATTAAGAACTGAAAGCGCAAGGTAAGAGCGCGCTGATTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTTGCTGCTACCGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTTGCTGCTACCGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTTAAAGACGTGGCGTGGATTAACGCTGCGCGAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTTAAAGACGTGGCGTGGATTAACGCTGCGCGAAGCG 2100
Qy 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAAACAAACACATGATGCGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAAACAAACACATGATGCGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCGCGTGG 2220
Db 2161 TCCATCGCAGAAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCGCGTGG 2220
Qy 2221 GCATGCTCAACATCGACACAGCAAGTGAATTTATGGTGAACGCTACTGCGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACAGCAAGTGAATTTATGGTGAACGCTACTGCGACCTTC 2280
Qy 2281 AAGGGTCAACCATCCAAACCGTTGGCTGCTGAGCGCAGGTATTAAACGCGCGAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTGGCTGCTGAGCGCAGGTATTAAACGCGCGAGT 2340
Qy 2341 CCAGAACAAAGAGCTGGCAAAAGATTTCTCGAAAACTATCTGCTGATGATGAAGTCTG 2400
Db 2341 CCAGAACAAAGAGCTGGCAAAAGATTTCTCGAAAACTATCTGCTGATGATGAAGTCTG 2400
Qy 2401 GAACGGTTAATAAAGACAAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAG 2460
Db 2401 GAACGGTTAATAAAGACAAAACCGCTGGGTGCGGTAGCGCTGAAAGTCTTACGAGGAAG 2460
Qy 2461 TTGGCGAAAGATCCACGATTTGCGCCACCATGGAAAAACCGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTTGCGCCACCATGGAAAAACCGCCAGAAAGGTGAAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAACGCC 2580
Qy 2581 GCGAGCGGTGCTGAGACTGCTGATGAGCCCTGAAAGACCGCGCAGACTTAATTCGAGCTCG 2640
Db 2581 GCGAGCGGTGCTGAGACTGCTGATGAGCCCTGAAAGACCGCGCAGACTTAATTCGAGCTCG 2640
Qy 2641 AACAAACAAACAAATAAACAATTAACAACCACTCGGATCGAGGAGGATTTTCAAGATTC 2700
Db 2641 AACAAACAAACAAATAAACAATTAACAACCACTCGGATCGAGGAGGATTTTCAAGATTC 2700
Qy 2701 GGATCTCTCTCTGTCGCGCGGCTCTGAGCGCGGCTCGAGCGCGGAGAACGCTTATGCTGT 2760
Db 2701 GGATCTCTCTCTGTCGCGCGGCTCTGAGCGCGGCTCGAGCGCGGAGAACGCTTATGCTGT 2760
Qy 2761 CCGGAATGTGGTAAGTTCCTTTCAGCGAGAGCTCTCCCTGCGCCACCGAGGTACCCAC 2820
Db 2761 CCGGAATGTGGTAAGTTCCTTTCAGCGAGAGCTCTCCCTGCGCCACCGAGGTACCCAC 2820

QY 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTCGCGCAAAATCTTTTAGCGAGTCCAGCAAC 2880
DB 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTCGCGCAAAATCTTTTAGCGAGCAGCAGC 2880
QY 2881 CTGGTGGCCATCAACGCACTCATCTGCGAGAGGCCATACAAATGTCAGAAATGTGC 2940
DB 2881 CTGGTGGCCATCAACGCACTCATCTGCGAGAGGCCATACAAATGTCAGAAATGTGC 2940
QY 2941 AAGTCTTTTCAGCGAGTCAGCAACCTGGTGGCCACCAACGTAATCTTTCAGCGAGAG 3000
DB 2941 AAGTCTTTTCAGCGAGTCAGCAACCTGGTGGCCACCAACGTAATCTTTCAGCGAGAG 3000
QY 3001 CCCTATGCTTGTCCGGAATGCTGTAAGTCTTTCAGCAGCAGAGTGCTCTTGTGTTAGACAC 3060
DB 3001 CCCTATGCTTGTCCGGAATGCTGTAAGTCTTTCAGCAGCAGAGTGCTCTTGTGTTAGACAC 3060
QY 3061 CAGCGTACCCACACGCGGTGAAAAACCGTATAAATGCCAGAGTCGCGCAAAATCTTTTAGC 3120
DB 3061 CAGCGTACCCACACGCGGTGAAAAACCGTATAAATGCCAGAGTCGCGCAAAATCTTTTAGC 3120
QY 3121 CAGCGCGGCCACTCGGAAGCCATCAACGCACTCATCTGCGAGAGGCCATACAAATGT 3180
DB 3121 CCGAGGGATACTGGTGGCCATCAACGCACTCATCTGCGAGAGGCCATACAAATGT 3180
QY 3181 CCAGAAATGGCAAGTCTTTCTCAACTTCAGCAACTTGGTCCGTCAACGTAATCTCAAC 3240
DB 3181 CCAGAAATGGCAAGTCTTTCTCAACTTCAGCAACTTGGTCCGTCAACGTAATCTCAAC 3240
QY 3241 ACCGCTAAAAAATAGTGGCCAGCGCCGAGTACCGGTAGCAGTTCGGAGTACCGCT 3300
DB 3241 ACCGCTAAAAAATAGTGGCCAGCGCCGAGTACCGGTAGCAGTTCGGAGTACCGCT 3300

RESULT 3

US-09-765-555-14
; Sequence 14, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; PRIOR FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m1 and zinc finger
; OTHER INFORMATION: protein ZFPm1

US-09-765-555-14

Query Match 97.8%; Score 3226.4; DB 10; Length 3300;
Best Local Similarity 98.6%; Pred. No. 0;
Matches 3254; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 1 CCGACACCATCAATGGTGCATAAACCTTTTCGGCGGTATGGCATGATAGCCCGGAAGAGA 60
DB 1 CCGACACCATCAATGGTGCATAAACCTTTTCGGCGGTATGGCATGATAGCCCGGAAGAGA 60
QY 61 GTCAATTACGGGTGGTGAATGTGAACACAGTAACGTTATACCATGTTCGAGAGTATGCCG 120
DB 61 GTCAATTACGGGTGGTGAATGTGAACACAGTAACGTTATACCATGTTCGAGAGTATGCCG 120
QY 121 GTGTCTTTATCAGACCGGTTTCCCGGTGGTGAACACGAGCCACCGCTTTCTCGCAAAA 180

DB 121 GTGTCTTTATCAGACCGGTTTCCCGGTGGTGAACACGAGCCAGCCACGCTTTCTCGAAAA 180
QY 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTACATTCCTCCAAACCGCGTGGCAC 240
DB 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTACATTCCTCCAAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAAAACAGTCTGTGATGGCGGTGGCACTTCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGGCAAAACAGTCTGTGATGGCGGTGGCACTTCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTCCCAATTTCTCGCGCGATTAATCTCGCGCGATCAATCTGGGTGGCCAGCG 360
DB 301 ACGCGCGTCCCAATTTCTCGCGCGATTAATCTCGCGCGATCAATCTGGGTGGCCAGCG 360
QY 361 TGGTGGTTCGATGGTAGAACGAGCGCGTTCGAAGCCCTGTAAAGCGCGGTGCAACAATC 420
DB 361 TGGTGGTTCGATGGTAGAACGAGCGCGTTCGAAGCCCTGTAAAGCGCGGTGCAACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGTGATCATTAATCTTCGCTGGATGACCAAGATGCA 480
DB 421 TTCTCGCGCAACCGCTCAGTGGGTGATCATTAATCTTCGCTGGATGACCAAGATGCA 480
QY 481 TTGCTGTGGAAGCTCCCTGCACTAATGTTCGCGCGTATTTCTTGATGCTCTGACCGA 540
DB 481 TTGCTGTGGAAGCTCCCTGCACTAATGTTCGCGCGTATTTCTTGATGCTCTGACCGA 540
QY 541 CACCCATCAACAGATTAATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
DB 541 CACCCATCAACAGATTAATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
QY 601 TGGTGGCATTTGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTCTGTCTCG 660
DB 601 TGGTGGCATTTGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTCTGTCTCG 660
QY 661 CGCGTCTCGCTCGCTGGCTGGCATTAATCTCACTCGCAATCAAAATCAGCCGATAG 720
DB 661 CGCGTCTCGCTCGCTGGCTGGCATTAATCTCACTCGCAATCAAAATCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
DB 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGGTGGCAACGATCAGATGGCGTGGCGCAA 840
DB 781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGGTGGCAACGATCAGATGGCGTGGCGCAA 840
QY 841 TGGCGGCCATTTACCGAGTCCCGGCTCGCGGTGGTGGCGGATATCTCGGTAGTGGGATAG 900
DB 841 TGGCGGCCATTTACCGAGTCCCGGCTCGCGGTGGTGGCGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATCAAAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGTCGCAACTCTCTCAGGCGCCAGCGGTGA 1020
DB 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGTGTCGCAACTCTCTCAGGCGCCAGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCGTCTCAGTGGTGAAGAAAAACCAACCTTCGGCGCCCAATA 1080
DB 1021 AGGGCAATCAGCTGTTGCCCGTCTCAGTGGTGAAGAAAAACCAACCTTCGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCTCCCGCGGTTGGCCGATTCATTAATGCAGTGGCAGCAGAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCTCCCGCGGTTGGCCGATTCATTAATGCAGTGGCAGCAGAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGGCAGTGCAGCGCAACCGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
DB 1141 CCCGACTGGAAGCGGGCAGTGCAGCGCAACCGCAATTAATGTAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAATTTCTCATGTTTTCAGAGCTTATCATCGACTGCAAGCGGTGCAACCAATGCTTCGGCG 1260

1201 GCACAAATCTCATGTTTGAAGCTTATCATCGACTGCAAGGTGCAACCAATGCTTCTGGCG 1260
1261 TCAGGAGGCAATCGGAGCTGTGGTATGCTGTGACAGTGTGTAATCACTCATATATTCG 1320
1261 TCAGGAGGCAATCGGAGCTGTGGTATGCTGTGACAGTGTGTAATCACTCATATATTCG 1320
1321 TGTGCTCAAGGGGCACTCCGTTCTGGATTAATGTTTTTGGCGGACATCAATACGGTT 1380
1321 TGTGCTCAAGGGGCACTCCGTTCTGGATTAATGTTTTTGGCGGACATCAATACGGTT 1380
1381 CTGGCAATATTTCTGAATAGCTGTGACAAATTAATCATCGCTCGTATATGTTGGA 1440
1381 CTGGCAATATTTCTGAATAGCTGTGACAAATTAATCATCGCTCGTATATGTTGGA 1440
1441 ATTGTAGCGGATAACAATTTTACACAGGAAACAGCCAGTCCGTTTGGTGTTCACGA 1500
1441 ATTGTAGCGGATAACAATTTTACACAGGAAACAGCCAGTCCGTTTGGTGTTCACGA 1500
1501 GCACCTCACCAACAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGTTAATCTGG 1560
1501 GCACCTCACCAACAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGTTAATCTGG 1560
1561 ATTAAACGGGATAAAGGCTATAACGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
1561 ATTAAACGGGATAAAGGCTATAACGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
1621 ACCGAAATTAAGTCAACCGTTGAGCATCCGATAAATCTGGAAGAGAAATTCGAGAGTT 1680
1621 ACCGAAATTAAGTCAACCGTTGAGCATCCGATAAATCTGGAAGAGAAATTCGAGAGTT 1680
1681 GGGCAATCGCGATGCGCTGACATTAATCTCTGGGACACGACCGCTTTGGTGGCTAC 1740
1681 GGGCAATCGCGATGCGCTGACATTAATCTCTGGGACACGACCGCTTTGGTGGCTAC 1740
1741 GCTCAATCTGGCTGTGGCTGAATCAACCGGACAAAGGTTTCCAGGACAAAGCTGAT 1800
1741 GCTCAATCTGGCTGTGGCTGAATCAACCGGACAAAGGTTTCCAGGACAAAGCTGAT 1800
1801 CCGTTTACCTGGATGCGCTGATTAACAAAGGCAAGCTGATGCTTACCCGATCGCTGTT 1860
1801 CCGTTTACCTGGATGCGCTGATTAACAAAGGCAAGCTGATGCTTACCCGATCGCTGTT 1860
1861 GAAGGTTTATCGCTGATTTAACAAGATCTGCTGCGAACCCGACAAACCTGGGAA 1920
1861 GAAGGTTTATCGCTGATTTAACAAGATCTGCTGCGAACCCGACAAACCTGGGAA 1920
1921 GAGATCCCGGCTGGATAAAGAACTGAAGAGGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
1921 GAGATCCCGGCTGGATAAAGAACTGAAGAGGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
1981 CTGCAAGAACCGTACTTCACTTGGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
1981 CTGCAAGAACCGTACTTCACTTGGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATTAAGCTGCGGCAAGAG 2100
2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATTAAGCTGCGGCAAGAG 2100
2101 GGTCTGACCTTCTCGTTGATGCTGATTAAGAAACAAACACATGAATGACAGACCGATTAAC 2160
2101 GGTCTGACCTTCTCGTTGATGCTGATTAAGAAACAAACACATGAATGACAGACCGATTAAC 2160
2161 TCCATCGCAGAGTCTGCTTTAATTAAGGCGAAACAGCGATGACCAATCAACCGGCGCTGG 2220
2161 TCCATCGCAGAGTCTGCTTTAATTAAGGCGAAACAGCGATGACCAATCAACCGGCGCTGG 2220
2221 GCATGGTCCAAATCGACACAGCAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
2221 GCATGGTCCAAATCGACACAGCAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
2281 AAGGTCACCAATCCAAACCGTCTGTTGGCTGCTGAGCGCAGGTATTAACCGGCGCAGT 2340
2281 AAGGTCACCAATCCAAACCGTCTGTTGGCTGCTGAGCGCAGGTATTAACCGGCGCAGT 2340

2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAAACTATCTGCTGACTGATGAAGGCTCG 2400
2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAAACTATCTGCTGACTGATGAAGGCTCG 2400
2401 GAAGCGGTTAATAAAGACAAACCGCTGGTGCCTAGCGCTGAAGTCTTACGAGGAAGAG 2460
2401 GAAGCGGTTAATAAAGACAAACCGCTGGTGCCTAGCGCTGAAGTCTTACGAGGAAGAG 2460
2461 TTGGCGAAAGATTCACGCTATTTGCGCCACATATGAAAAACGCCAGAAAGGTGAAATCATG 2520
2461 TTGGCGAAAGATTCACGCTATTTGCGCCACATATGAAAAACGCCAGAAAGGTGAAATCATG 2520
2521 CCGAACATCCGCGAGATGCTGGCTTTCTGCTGATGCGCTGCTGCTGCTGCTGCTGCTG 2580
2521 CCGAACATCCGCGAGATGCTGGCTTTCTGCTGATGCGCTGCTGCTGCTGCTGCTGCTG 2580
2581 GCCAGCGGTGCTCAGACTGCTGATGAAGCCCTGAAAGACCGCAGACGCTAAATTCGAGCTCG 2640
2581 GCCAGCGGTGCTCAGACTGCTGATGAAGCCCTGAAAGACCGCAGACGCTAAATTCGAGCTCG 2640
2641 AACCAACAAACAAATTAACAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
2641 AACCAACAAACAAATTAACAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
2701 GGATCTCTTCTCTGCTGGCCAGGGGCTCTGAGCCCGGGGAGAGCCCTATGCTTGT 2760
2701 GGATCTCTTCTCTGCTGGCCAGGGGCTCTGAGCCCGGGGAGAGCCCTATGCTTGT 2760
2761 CCGGAATGTGTTAAGTCTCTTTCAGCCAGAGCAGCTCCTGCTGGCGCCACAGCCTGATCCAC 2820
2761 CCGGAATGTGTTAAGTCTCTTTCAGCCAGAGCTCTCCTGCTGGCGCCACAGCCTGATCCAC 2820
2821 ACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGGCAAAATCTTTTACCGAGTCCAGCAAC 2880
2821 ACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGGCAAAATCTTTTACCGAGTCCAGCAAC 2880
2881 CTGCTGGCCATCAACGCACTCATCTGCGAGAGCGCATACAAATGTCAGAAATGTTGGC 2940
2881 CTGCTGGCCATCAACGCACTCATCTGCGAGAGCGCATACAAATGTCAGAAATGTTGGC 2940
2941 AAGTCTTTTACCGAGTCCAGCAACCTGCTGGCGCCACCAACGTAATCAACCGGGGAGAAG 3000
2941 AAGTCTTTTCTCGGCTGACAACTCTGCTGGCGCAACCAACGTAATCAACCGGGGAGAAG 3000
3001 CCTATGCTGCTGGCAATGTTGTAAGTCTTTCAGCAACCGAGTGGCTCTTGGTTAGCAC 3060
3001 CCTATGCTGCTGGCAATGTTGTAAGTCTTTCAGCAACCGAGTGGCTCTTGGTTAGCAC 3060
3061 CAGCGTACCCACACCGGGTGAATAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTTAGC 3120
3061 CAGCGTACCCACACCGGGTGAATAACCGTATAAATGCCAGAGTGGCGCAAACTCTTTTAGC 3120
3121 CAGCGCGCCACCTGGAACCGCATCAACGCACTCATCTGCGGAGAGGCAATCAACAATGT 3180
3121 CAGCGCGCCACCTGCGGAGCCATCAACGCACTCATCTGCGGAGAGGCAATCAACAATGT 3180
3181 CCAGAATGTGGCAAGTCTTTTCAACTTCAGGCAACTCTGCTGGTCCGTCACCAAGTACTCAC 3240
3181 CCAGAATGTGGCAAGTCTTTTCAACTTCAGGCAACTCTGCTGGTCCGTCACCAAGTACTCAC 3240
3241 ACCGGTAAAAAACTAGTGGCGCGCGCCAGTACCCGTAAGAGCGTTCCGGACTACGCT 3300
3241 ACCGGTAAAAAACTAGTGGCGCGCGCCAGTACCCGTAAGAGCGTTCCGGACTACGCT 3300

RESULT 4
US-09-765-555-15
; Sequence 15, Application US/09765555
; Publication No. US2003003735A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate

; TITLE OF INVENTION: expression in plants	
; FILE REFERENCE: 27801-20014.40	
; CURRENT APPLICATION NUMBER: US/09/765,555	
; CURRENT FILING DATE: 2002-05-24	
; PRIOR APPLICATION NUMBER: US 09/620,897	
; PRIOR FILING DATE: 2000-01-21	
; PRIOR APPLICATION NUMBER: US 60/177,468	
; PRIOR FILING DATE: 2000-01-21	
; NUMBER OF SEQ ID NOS: 75	
; SOFTWARE: FastSeq for Windows Version 4.0	
; SEQ ID NO 15	
; LENGTH: 3300	
; TYPE: DNA	
; ORGANISM: Artificial Sequence	
; FEATURE:	
; OTHER INFORMATION: Partial sequence of pMal-m2 and zinc finger	
; OTHER INFORMATION: protein 2FPm2	
US-09-765-555-15	
Query Match 97.8%; Score 3226.4; DB 10; Length 3300;	
Best Local Similarity 98.6%; Pred. No. 0;	
Matches 3254; Conservative 0; Mismatches 46; Indels 0; Gaps 0;	
QY	1 CGGACACCATCGAATGGTGCAGAACTTTTCGGGTATGCGATAGAGCGCCCGGAGAGA 60
DB	1 CGGACACCATCGAATGGTGCAGAACTTTTCGGGTATGCGATAGAGCGCCCGGAGAGA 60
QY	61 GTCAATTACAGGTGGTGAATGTGAACACCAAGTAACGTTATACGATGTCGACAGTATGCGG 120
DB	61 GTCAATTACAGGTGGTGAATGTGAACACCAAGTAACGTTATACGATGTCGACAGTATGCGG 120
QY	121 GTGTCTTTATCAGACCGTTTCCGGGTGGTGAACACGAGCCAGCCACGTTTCTGCGAAAA 180
DB	121 GTGTCTTTATCAGACCGTTTCCGGGTGGTGAACACGAGCCAGCCACGTTTCTGCGAAAA 180
QY	181 CCGCGGAAAGTGAAGCGGCGATGGCGAGCTGAATTTACATTTCCACACCGGTGGCAC 240
DB	181 CCGCGGAAAGTGAAGCGGCGATGGCGAGCTGAATTTACATTTCCACACCGGTGGCAC 240
QY	241 AACAACTGCGCGGCAAAACAGTGTGCTGATTTGGGTTTCCACCTCCAGTCTGGCCCTGC 300
DB	241 AACAACTGCGCGGCAAAACAGTGTGCTGATTTGGGTTTCCACCTCCAGTCTGGCCCTGC 300
QY	301 ACGCGCGTTCGCAATTTGTCGGCGGATTAATCTCGGCCGATCAACTGGGTGCCAGCG 360
DB	301 ACGCGCGTTCGCAATTTGTCGGCGGATTAATCTCGGCCGATCAACTGGGTGCCAGCG 360
QY	361 TGGTGGTTCGATGTAGAACGAGCGGCTCGAAGCCTGTAAAGCGGCGTGCACAAATC 420
DB	361 TGGTGGTTCGATGTAGAACGAGCGGCTCGAAGCCTGTAAAGCGGCGTGCACAAATC 420
QY	421 TTCTCGCGCAACCGCTCAGTGGCTGATCATTTAACTATCCGCTGGATGACAGGATGCCA 480
DB	421 TTCTCGCGCAACCGCTCAGTGGCTGATCATTTAACTATCCGCTGGATGACAGGATGCCA 480
QY	481 TTGCTGTGGAAGCTGCTGCATTAATGTTCCGGCGTTATTTCTGTGATGTTCTGACACGA 540
DB	481 TTGCTGTGGAAGCTGCTGCATTAATGTTCCGGCGTTA TTCTGTGATGTTCTGACACGA 540
QY	541 CACCCATCAACAGTATTATTTTCTCCCATGAAGAAGGATCGGATCGGCGTGGAGCATC 600
DB	541 CACCCATCAACAGTATTATTTTCTCCCATGAAGAAGGATCGGATCGGCGTGGAGCATC 600
QY	601 TGGTGCATTTGGTTCACAGCAAACTCGCGCTGTAGCGGCGCCATTAAGTTCGTCTCGG 660
DB	601 TGGTGCATTTGGTTCACAGCAAACTCGCGCTGTAGCGGCGCCATTAAGTTCGTCTCGG 660
QY	661 CCGGCTCGGCTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB	661 CCGGCTCTGGCTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY	721 CCGAACGGGAAGCGCATCGAGTGCATGTCGGGTTTTCAAACCAACCATGCAAAATGCTGA 780
DB	721 CCGAACGGGAAGCGCATCGAGTGCATGTCGGGTTTTCAAACCAACCATGCAAAATGCTGA 780

DB	721 CGGAACGGGAAGCGCATCGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
QY	781 ATGAGGGCATCTGTTCCCATCGCATGCTGGTGTGCGCAACGATCAGATGCGGTGGCGGCAA 840
DB	781 ATGAGGGCATCTGTTCCCATCGCATGCTGGTGTGCGCAACGATCAGATGCGGTGGCGGCAA 840
QY	841 TGGCGGCATTTACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGTAGTGGGATAGG 900
DB	841 TGGCGGCATTTACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGTAGTGGGATAGG 900
QY	901 ACGATACCGGAAGACAGCTCATGTTTATATCCCGCCGCTTAACACCATCAAAACAGGATTTTC 960
DB	901 ACGATACCGGAAGACAGCTCATGTTTATATCCCGCCGCTTAACACCATCAAAACAGGATTTTC 960
QY	961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCCACGCGGTGA 1020
DB	961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCCACGCGGTGA 1020
QY	1021 AGGGCAATCAGCTGTTGCCCGCTCTCACTGTTGAAAGAAACCAACCTGGCGCCCAATA 1080
DB	1021 AGGGCAATCAGCTGTTGCCCGCTCTCACTGTTGAAAGAAACCAACCTGGCGCCCAATA 1080
QY	1081 CGCAAAACCGCTCTCTCCCGCGCTTGGCCGATTCAATTAATGCACTGGCAGCAGAGTTT 1140
DB	1081 CGCAAAACCGCTCTCTCCCGCGCTTGGCCGATTCAATTAATGCACTGGCAGCAGAGTTT 1140
QY	1141 CCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
DB	1141 CCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
QY	1201 GCACAAATTTCTATGTTTGACAGCTTATCATCGATGCGAGGTGCAACCAATGTTCTGGCG 1260
DB	1201 GCACAAATTTCTATGTTTGACAGCTTATCATCGATGCGAGGTGCAACCAATGTTCTGGCG 1260
QY	1261 TCAGGACGCAATCGGAAGCTGTGGTATGCTGTGAGGTGCGTAAATCACTGCATAAATTCG 1320
DB	1261 TCAGGACGCAATCGGAAGCTGTGGTATGCTGTGAGGTGCGTAAATCACTGCATAAATTCG 1320
QY	1321 TGTGCTCAAGCGCACTCCCGTTCGATTAATGTTTGTGGCGCAATCAATAACGGTT 1380
DB	1321 TGTGCTCAAGCGCACTCCCGTTCGATTAATGTTTGTGGCGCAATCAATAACGGTT 1380
QY	1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
DB	1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
QY	1441 ATTGTGAGCGGATACAAATTTTACACAGCAACACGAGTCCGTTTAGGTGTTTTCACGA 1500
DB	1441 ATTGTGAGCGGATACAAATTTTACACAGCAACACGAGTCCGTTTAGGTGTTTTCACGA 1500
QY	1501 GCACCTTCCAAACAAAGGACCATAGATTATGAAAACCTGAAGAGGTAAACTGGTAATCTGG 1560
DB	1501 GCACCTTCCAAACAAAGGACCATAGATTATGAAAACCTGAAGAGGTAAACTGGTAATCTGG 1560
QY	1561 ATTAAACGGCGATAAAGGCTATAACCGCTCTCGCTGAAGTCCGTTAAGAAATTCAGAAAGAT 1620
DB	1561 ATTAAACGGCGATAAAGGCTATAACCGCTCTCGCTGAAGTCCGTTAAGAAATTCAGAAAGAT 1620
QY	1621 ACCGGAATTTAAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGGTT 1680
DB	1621 ACCGGAATTTAAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACACAGGTT 1680
QY	1681 GCGGCAACTGGCGATGGCCCTGACATTTCTTCTGGGCAACACGACCGCTTTGGTGGCTAC 1740
DB	1681 GCGGCAACTGGCGATGGCCCTGACATTTCTTCTGGGCAACACGACCGCTTTGGTGGCTAC 1740
QY	1741 GCTCAATCTGGCTGTTGGCTGAAATCAACCCCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
DB	1741 GCTCAATCTGGCTGTTGGCTGAAATCAACCCCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
QY	1801 CCGTTTACCTGGATGCGGTGATCAACGCGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
DB	1801 CCGTTTACCTGGATGCGGTGATCAACGCGCAAGCTGATTGCTTACCCGATCGCTGTT 1860

1861 GAAGGTTATCGTGATTTTATCAAGATCTGTCGCAACCCGCCAAAAACCTGGGAA 1920
1861 GAAGGTTATCGTGATTTTATCAAGATCTGTCGCAACCCGCCAAAAACCTGGGAA 1920
1921 GAGATCCCGCGCTGGATGAAGAACTGAAGCGAAGAGTGAAGCGCTGATGTTCAAC 1980
1921 GAGATCCCGCGCTGGATGAAGAACTGAAGCGAAGAGTGAAGCGCTGATGTTCAAC 1980
1981 CTGCAAGAACCGTACTTCACTTGGCGCTGATTGCTGTCGAGCGGGGTTATGCGTTCAAG 2040
1981 CTGCAAGAACCGTACTTCACTTGGCGCTGATTGCTGTCGAGCGGGGTTATGCGTTCAAG 2040
2041 TATGAAAAACCGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGGAAAGCG 2100
2041 TATGAAAAACCGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGGAAAGCG 2100
2101 GGTCTGACCTTCTCGTGGTGAACCTGATTAAACCAACACATGAATGCGACACACCGATTAC 2160
2101 GGTCTGACCTTCTCGTGGTGAACCTGATTAAACCAACACATGAATGCGACACACCGATTAC 2160
2161 TCCATCGCAGAGCTGCTTTTAAAGCGCAACAGCGATGACCATCAACGGCCGCTGG 2220
2161 TCCATCGCAGAGCTGCTTTTAAAGCGCAACAGCGATGACCATCAACGGCCGCTGG 2220
2221 GCATGGTCAACATCGACACACGAAAGTGAATTATGTTAAACGGTACTGCCGACCTTC 2280
2221 GCATGGTCAACATCGACACACGAAAGTGAATTATGTTAAACGGTACTGCCGACCTTC 2280
2281 AAGGTCACCAATCAACACCGTTCGTCGCTGTCGAGCGAGGTATTAACGCGCCAGT 2340
2281 AAGGTCACCAATCAACACCGTTCGTCGCTGTCGAGCGAGGTATTAACGCGCCAGT 2340
2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGCTG 2400
2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGCTG 2400
2401 GAAGCGGTTAATGAAGCAACCGTGGTGGCTGAGCGTGAAGTCTTACGAGGAAGAG 2460
2401 GAAGCGGTTAATGAAGCAACCGTGGTGGCTGAGCGTGAAGTCTTACGAGGAAGAG 2460
2461 TTGGCAAGATCCACGATTTGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
2461 TTGGCAAGATCCACGATTTGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
2521 CCGAACATCCGCGAGATGTCGCTTTCTGTTATGCGTATGCGGTGATCAACGCC 2580
2521 CCGAACATCCGCGAGATGTCGCTTTCTGTTATGCGTATGCGGTGATCAACGCC 2580
2581 GCCAGCGTCTGACGATGTCGATGAGCCCTGGAAGACGCGGAGACTAATTCGAGCTCG 2640
2581 GCCAGCGTCTGACGATGTCGATGAGCCCTGGAAGACGCGGAGACTAATTCGAGCTCG 2640
2641 AACCAACAAACAATAACAATAACAACCTCGGGATCGAGGGAAGATTTACAGATTTC 2700
2641 AACCAACAAACAATAACAATAACAACCTCGGGATCGAGGGAAGATTTACAGATTTC 2700
2701 GGATCTCTTCTCTGTCGCGCCAGCGGCTCTCGAGCCCGGGGAGAGCCCTATGTTGT 2760
2701 GGATCTCTTCTCTGTCGCGCCAGCGGCTCTCGAGCCCGGGGAGAGCCCTATGTTGT 2760
2761 CCGGAATGTGGTAAAGTCTCTTCAGCGCAGAGCTCCCTGGTGGCCACGAGCTACCCAC 2820
2761 CCGGAATGTGGTAAAGTCTCTTCAGAGCTCTCACCTGGTGGCCACGAGCTACCCAC 2820
2821 ACGGGTGAAAAACCGTATAATGCCAGAGTGGGCAAACTCTTTAGCCAGTCCAGCAAC 2880
2821 ACGGGTGAAAAACCGTATAATGCCAGAGTGGGCAAACTCTTTAGCCAGTCCAGCAAC 2880
2881 CTGCTGCGCATCAACGCTACTTACGAGAGGAGCCATCAAAATGTCAGAAATGTGGC 2940
2881 CTGCTGCGCATCAACGCTACTTACGAGAGGAGCCATCAAAATGTCAGAAATGTGGC 2940

2941 AAGTCTTTACGCCAGTCCAGCAACCTGGTGGCGCCACCAACGTACTCAACCGGGAGAG 3000
2941 AAGTCTTTACGCCAGTCCAGCAACCTGGTGGCGCCACCAACGTACTCAACCGGGAGAG 3000
3001 CCCTATGCTTGTCCGGAATGTGGTAAAGTCTTTCAGCACCAGTGGCTCCTTGGTTAGCAC 3060
3001 CCCTATGCTTGTCCGGAATGTGGTAAAGTCTTTCAGCACCAGTGGCTCCTTGGTTAGCAC 3060
3061 CAGGTACCCACACGCGGTGAAAAACCGTATAAAATGCCAGAGTGGCGCAAAATCTTTAGC 3120
3061 CAGGTACCCACACGCGGTGAAAAACCGTATAAAATGCCAGAGTGGCGCAAAATCTTTAGC 3120
3121 CAGCGCGCCACCTGGAAACGCAATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
3121 CAGCGCGCCACCTGGCGGAGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
3181 CCAGAAATGTGGCAAGTCTTTCTCAACTTCAGGCAACTTGGTCCGTCAACCAACGTACTCAC 3240
3181 CCAGAAATGTGGCAAGTCTTTCTCTCGGTCTGACAATCTCGTCCGGCACCAACGTACTCAC 3240
3241 ACCGTTAAAAAACTAGTGGCCAGCGCCGCGCAGTACCCGTACGACGTTCCGGAATCGCT 3300
3241 ACCGTTAAAAAACTAGTGGCCAGCGCCGCGCAGTACCCGTACGACGTTCCGGAATCGCT 3300

RESULT 5
US-09-765-555-16
; Sequence 16, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE REFERENCE: 27801-28014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m3 and zinc finger
; OTHER INFORMATION: protein ZFPm3
US-09-765-555-16

Query Match 97.5%; Score 3216.8; DB 10; Length 3300;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 3248; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGCATAAACCTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
DB 1 CCGACACCATCGAATGGTGCATAAACCTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTCAGGGTGGTGAATGTGAACCAAGTAACTGATGATGTCGAGAGTATGCCG 120
DB 61 GTCAATTCAGGGTGGTGAATGTGAACCAAGTAACTGATGATGTCGAGAGTATGCCG 120
QY 121 GTGTCCTTTATCAGACCGCTTTCGGGTGGTGAACCAAGTAACTGATGATGTCGAGAGTATGCCG 180
DB 121 GTGTCCTTTATCAGACCGCTTTCGGGTGGTGAACCAAGTAACTGATGATGTCGAGAGTATGCCG 180
QY 181 CCGCGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
DB 181 CCGCGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
QY 241 AACCACTGGCGGCAACACAGTGTGCTGATTTGGGTGGCTTCCACCTCCAGTCTGGCCCTGC 300

QY	2461	TTGGCGAAAGATCCACGCTATTGGCGCCACCAATGGAAGACCCCGAAGAGTGAAATCATG	2520
Db	2461	TTGGCGAAAGATCCACGCTATTGGCGCCACCAATGGAAGACCCCGAAGAGTGAAATCATG	2520
QY	2521	CCGAACATCCCGCAGATGTCGCGCTTCTGGTGATCCGCTGCGTACTGCGGTGATCAACGCC	2580
Db	2521	CCGAACATCCCGCAGATGTCGCGCTTCTGGTGATCCGCTGCGTACTGCGGTGATCAACGCC	2580
QY	2581	GCCAGCGTCTGCAGACTGTGCATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG	2640
Db	2581	GCCAGCGTCTGCAGACTGTGCATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG	2640
QY	2641	AACACACACACAAATACAAATAACAAACCTCCGGATCGAGGGAAGATTTTCAGAAATTC	2700
Db	2641	AACACACACACAAATACAAATAACAAACCTCCGGATCGAGGGAAGATTTTCAGAAATTC	2700
QY	2701	GGATCTCTTCTCTGTGGCCAGCGCGCCCTCGAGCCCGGGGAGAAGCCCTATGCTTGT	2760
Db	2701	GGATCTCTTCTCTGTGGCCAGCGCGCCCTCGAGCCCGGGGAGAAGCCCTATGCTTGT	2760
QY	2761	CCGGAATGTGGTAAGTCTTTCAGCAGAGCAGCTCCTCTGTGTGGCCACCAAGCGTACCAC	2820
Db	2761	CCGGAATGTGGTAAGTCTTTCAGCAGAGCAGCTCCTCTGTGTGGCCACCAAGCGTACCAC	2820
QY	2821	ACGGGTGAAGAAACCGTATAATGCCAGAGTGGCGCAAACTCTTTAGCCAGTCCAGCAAC	2880
Db	2821	ACGGGTGAAGAAACCGTATAATGCCAGAGTGGCGCAAACTCTTTAGCCAGTCCAGCAAC	2880
QY	2881	CTGTGCGCCATCAACGACCTCATCTATCTGGGAGAAGCCATACAAATGTCAGAAATGTGGC	2940
Db	2881	CTGTGCGCCATCAACGACCTCATCTATCTGGGAGAAGCCATACAAATGTCAGAAATGTGGC	2940
QY	2941	AAGTCTTTACGCCAGTCCAGCAACCTGTGGCGCCACCAAGCTACTCACCCGGGAGAAG	3000
Db	2941	AAGTCTTTACGCCAGTCCAGCAACCTGTGGCGCCACCAAGCTACTCACCCGGGAGAAG	3000
QY	3001	CCCTATGCTTGTCCGGAATGTGGTAAGTCTTTCAGCACCAAGTGGCTCTTGTGTAGACAC	3060
Db	3001	CCCTATGCTTGTCCGGAATGTGGTAAGTCTTTCAGCACCAAGTGGCTCTTGTGTAGACAC	3060
QY	3061	CAGGTAACCCACACGGGTGAAAAACCGTATAATGCCAGAGTGGCGCAAACTTTTATAGC	3120
Db	3061	CAGGTAACCCACACGGGTGAAAAACCGTATAATGCCAGAGTGGCGCAAACTTTTATAGC	3120
QY	3121	CAGCGCGCCACCTGTGAACGCCATCAACGACCTCATCTGGGAGAAGCCATACAAATGT	3180
Db	3121	GACTGCGGACCTTGCTCGCCATCAAGCACTCATCTGGGAGAAGCCATACAAATGT	3180
QY	3181	CCAGATGTGGCAAGTCTTTCTCACTTAGGCAACTTGGTTCGCTCACCAAGCTACTCAC	3240
Db	3181	CCAGATGTGGCAAGTCTTTCTCACTTAGGCAACTTGGTTCGCTCACCAAGCTACTCAC	3240
QY	3241	ACCGGTAAAAAACTAGTGGCCAGGCGGCGCAGTACCCTGACCAAGCTCCGGACTACGCT	3300
Db	3241	ACCGGTAAAAAACTAGTGGCCAGGCGGCGCAGTACCCTGACCAAGCTCCGGACTACGCT	3300
RESULT 6			
US-10-257-384A-1			
; Sequence 1, Application US/10257384A			
; Publication No. US20040087524A1			
; GENERAL INFORMATION:			
; APPLICANT: Wiedersanders, Bernd			
; APPLICANT: Maubach, Gunter			
; TITLE OF INVENTION: Agent for postoperative use after removal of bone tumors			
; FILE REFERENCE: 2945-101			
; CURRENT APPLICATION NUMBER: US/10/257,384A			
; CURRENT FILING DATE: 2003-08-21			
; PRIOR APPLICATION NUMBER: PCT/DE 01/01510			
; PRIOR FILING DATE: 2001-04-18			
; PRIOR APPLICATION NUMBER: DE 100 20 125.3			
; PRIOR FILING DATE: 2000-04-18			
; NUMBER OF SEQ ID NOS: 19			

Db 2701 GGATCCTCCAGTCCCGCAAGCCGCGCGCTGGTGGGAGG 2741

RESULT 7
US-10-149-472-5
; Sequence 5, Application US/10149472
; Publication No. US20040029204A1
; GENERAL INFORMATION:
; APPLICANT: GOUBIN-GRAMATICA, FRANCOISE
; APPLICANT: DUCOMMUN, BERNARD
; APPLICANT: PREVOST, GREGOIRE
; TITLE OF INVENTION: METHOD FOR OBTAINING HUMAN CDC25 PHOSPHATASES AND METHOD
; FILE REFERENCE: bml-427.065
; CURRENT APPLICATION NUMBER: US/10/149,472
; PRIOR FILING DATE: 2002-06-07
; PRIOR FILING DATE: 2000-12-13
; PRIOR APPLICATION NUMBER: FR 99/15722
; PRIOR FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: FR 00/06883
; PRIOR FILING DATE: 2000-05-30
; PRIOR APPLICATION NUMBER: FR 00/12008
; PRIOR FILING DATE: 2000-09-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 8101
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: JM109 / pMAL - Hs Cdc35C strain
US-10-149-472-5

Query Match 82.0%; Score 2705.8; DB 17; Length 8101;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CGACACCATCGAATGGTGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 1 CGACACCATCGAATGGTGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60

Qy 61 GTCGAATTCAGGGTGGTGAATGTGAACCAAGTAACGTTATACGATGTGCGAGAGTATGCCG 120
Db 61 GTCGAATTCAGGGTGGTGAATGTGAACCAAGTAACGTTATACGATGTGCGAGAGTATGCCG 120

Qy 121 GTGTCTCTTATCAGACCGGTTTCCCGCGTGGTGAACCAAGCGGCGAGCCAGCTTCTCGCAAAA 180
Db 121 GTGTCTCTTATCAGACCGGTTTCCCGCGTGGTGAACCAAGCGGCGAGCCAGCTTCTCGCAAAA 180

Qy 181 CGCGGGAAGAAAGTGGAGCGCGATGGCGGAGCTGAATTAATTCACCAACCGCGTGGCAC 240
Db 181 CGCGGGAAGAAAGTGGAGCGCGATGGCGGAGCTGAATTAATTCACCAACCGCGTGGCAC 240

Qy 241 AACAACTGGCGGCAACAGTCGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGCCCTGCG 300
Db 241 AACAACTGGCGGCAACAGTCGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGCCCTGCG 300

Qy 301 ACGCGCGTGCAGAAATGTGCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGTGCAGAAATGTGCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360

Qy 361 TGGTGGTGTGATGGTGAAGCAAGCGCGTGCAGAGCTGTAAAGCGCGGCTGCACAATC 420
Db 361 TGGTGGTGTGATGGTGAAGCAAGCGCGTGCAGAGCTGTAAAGCGCGGCTGCACAATC 420

Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATTAATCTCGCTGGATGACCGAGTGCCTCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATTAATCTCGCTGGATGACCGAGTGCCTCA 480

Qy 481 TTGCTGTGGAAGCTGCCTGCATAATGTTCCGGGTTATTTCTTGATGTTCTTGACGAGA 540
Db 481 TTGCTGTGGAAGCTGCCTGCATAATGTTCCGGGTTATTTCTTGATGTTCTTGACGAGA 540

Db 481 TTGCTGTGGAAGCTGCCTGCATAATGTTCCGGGTTATTTCTTGATGTTCTTGACGAGA 540
Qy 541 CACCCATCAACAGTATTATTTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTCCGATTTGGGTACCAGCAATCGCGTGTAGCGGGCCATTAAAGTTCCTCTCGG 660
Db 601 TGGTCCGATTTGGGTACCAGCAATCGCGTGTAGCGGGCCATTAAAGTTCCTCTCGG 660
Qy 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGCGACTGGAAGTCCCATGTCGGTTTTTCAACAAACCATCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGCGACTGGAAGTCCCATGTCGGTTTTTCAACAAACCATCAAAATGCTGA 780
Qy 781 ATGAGGGCATCGTTTCCCACTGGCATGCTGTTGCCAAGATCAGATGGCGCTGGCGCAA 840
Db 781 ATGAGGGCATCGTTTCCCACTGGCATGCTGTTGCCAAGATCAGATGGCGCTGGCGCAA 840
Qy 841 TGGCGCCATTACCGAGTCCGGGCTGGCGGCTGGCGGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGCCATTACCGAGTCCGGGCTGGCGGCTGGCGGATATCTCGGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAACCATCAAAATGCTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCAACCATCAAAATGCTTC 960
Qy 961 GCCTGCTGGGGCAAAACCGAGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGCTGGGGCAAAACCGAGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTGCGGCTCTCACTGGTGAAGAAAAACACCTGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGCGGCTCTCACTGGTGAAGAAAAACACCTGCGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGGCTGGCGGATTCATTATGCACTGGCAAGCAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGCTGGCGGATTCATTATGCACTGGCAAGCAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTGAGTTAGTCTCACTATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTGAGTTAGTCTCACTATTAG 1200
Qy 1201 GCACAAATTCATGTTTGAAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATTCATGTTTGAAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGCGAGCATCGGAAGCTGTGATGCTGTGCAAGTCTGTAATCACTGCAATAATTCG 1320
Db 1261 TCAGGCGAGCATCGGAAGCTGTGATGCTGTGCAAGTCTGTAATCACTGCAATAATTCG 1320
Qy 1321 TGTGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAACCGTT 1380
Db 1321 TGTGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTGGCGCGACATCAACCGTT 1380
Qy 1381 CTGGCAAAATTCATGAAATGAGCTGTGCAATTAATCATCGGCTGTAATGTTGGA 1440
Db 1381 CTGGCAAAATTCATGAAATGAGCTGTGCAATTAATCATCGGCTGTAATGTTGGA 1440
Qy 1441 ATTGTAGCGGATAACAAATTTTCCAGGAAAAACAGCAGTCCGTTTAGGTGTTTTTCAACA 1500
Db 1441 ATTGTAGCGGATAACAAATTTTCCAGGAAAAACAGCAGTCCGTTTAGGTGTTTTTCAACA 1500
Qy 1501 GCACCTTCAACCAAGGACCATAGATTATGAAGAACTGAAGAGGTAAACTGGTAATCTCG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTATGAAGAACTGAAGAGGTAAACTGGTAATCTCG 1560
Qy 1561 ATTAAACGGCGATAAAGGCTTAAACCGTCTCGCTGAAGTGGTAAGAAATTCAGGAAGAT 1620
Db 1561 ATTAAACGGCGATAAAGGCTTAAACCGTCTCGCTGAAGTGGTAAGAAATTCAGGAAGAT 1620

Db 541 CACCCATCAACAGTATTATTTCTCCATGAAGACGGTACGCGACTGGCGCTGGAGCATC 600
Qy 601 TGGTCGCAATTGGGTACACGAAATCGCGCTGTGTAGCGGCCCATTTAACTGTTCTGCTCGG 660
Db 601 TGGTCGCAATTGGGTACACGAAATCGCGCTGTGTAGCGGCCCATTTAACTGTTCTGCTCGG 660
Qy 661 CGGCTCTGGCTCTGGCTGGCTGGCTGAATAATATCTCATCTCGCAATCAAAATTCAGCGGATAG 720
Db 661 CGGCTCTGGCTCTGGCTGGCTGGCTGAATAATATCTCATCTCGCAATCAAAATTCAGCGGATAG 720
Qy 721 CGGAACGGGAACGGCACTGGAGTGCATGTCGGTTTCAAAACCAATGCAAAATGCTGA 780
Db 721 CGGAACGGGAACGGCACTGGAGTGCATGTCGGTTTCAAAACCAATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATCTGTTCCCACTGGGATGCTGGTTGCCAAACGATCAGATGGCGCTGGCGGCAA 840
Db 781 ATGAGGGCATCTGTTCCCACTGGGATGCTGGTTGCCAAACGATCAGATGGCGCTGGCGGCAA 840
Qy 841 TGGCGGCCATTAACCGAGTCCGGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGGCCATTAACCGAGTCCGGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Qy 901 AGGATACCGGAACAGACAGCTCATGTTATATCCCGCGTTAACCAACATCAAAACAGATTTTC 960
Db 901 AGGATACCGGAACAGACAGCTCATGTTATATCCCGCGTTAACCAACATCAAAACAGATTTTC 960
Qy 961 GCCTGCTGGGCAAAACCGAGCTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAAAACCGAGCTGGAGCCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTTGCCGCTCTCATCTGGTGAAGAAAGAAACCAACCTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCGCTCTCATCTGGTGAAGAAAGAAACCAACCTGGCGGCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGGTTGGCCGATTCAATTAATGCAGTGGCAAGAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCCGATTCAATTAATGCAGTGGCAAGAGGTTT 1140
Qy 1141 CCGGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Db 1141 CCGGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTTAGCTCACTCATTTAG 1200
Qy 1201 GCACAATTTCTATGTTTCAGAGCTTATCATCGACTGACGGTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTTCTATGTTTTCAGAGCTTATCATCGACTGACGGTGACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGCGAGTTCGTAATCACTGCATAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGCGAGTTCGTAATCACTGCATAATTCG 1320
Qy 1321 TGTGCTCAAGCGGCATCTCCGTTCTGGATTAATGTTTTTGGCGGCAATCATAACCGTT 1380
Db 1321 TGTGCTCAAGCGGCATCTCCGTTCTGGATTAATGTTTTTGGCGGCAATCATAACCGTT 1380
Qy 1381 CTGGCAATAATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA 1440
Db 1381 CTGGCAATAATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGTGGA 1440
Qy 1441 ATTGTGAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAGTGTGTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAGTGTGTTTTCACGA 1500
Qy 1501 GCATTCACCAACAGGACCAATAGATTATGAAAACCTGAAGAGTAACTGGTAATCTGG 1560
Db 1501 GCATTCACCAACAGGACCAATAGATATGAAAATCGAAGAGTAACTGGTAATCTGG 1560
Qy 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCCGTAAGAAATTCGGAAGAT 1620
Db 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCCGTAAGAAATTCGGAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAACTGGAAGAGAAATTTCCCAACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAACTGGAAGAGAAATTTCCCAACAGGTT 1680

Qy 1681 GCGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGCAAAAGGTTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGCAAAAGGTTTCCAGGACAAGCTGTAT 1800
Qy 1801 CCGTTTACTCTGGGATGGCGTACGTTTACAACGGCAAGCTGATTGCTTACC CGGATCGCTGTT 1860
Db 1801 CCGTTTACTCTGGGATGGCGTACGTTTACAACGGCAAGCTGATTGCTTACC CGGATCGCTGTT 1860
Qy 1861 GAACGGTTATCGCTGATTATACAAAGATCTGCTGCGCAACCCCGCAAAACCTGGGAA 1920
Db 1861 GAACGGTTATCGCTGATTATACAAAGATCTGCTGCGCAACCCCGCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGGAAAGCTAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAAACGTTACTTCACTGCGCCGCTGATTGCTGCTGACCGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAAACGTTACTTCACTGCGCCGCTGATTGCTGCTGACCGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTAAGACGTTGAGGCTGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGTTGAGGCTGATTAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAATAAAGGCGAAACAGCGATGACCGATTAAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTAAATAAAGGCGAAACAGCGATGACCGATTAAC 2160
Qy 2161 TCCATCGCAGAAGCTGCTTTAATAAGGCGAAACAGCGATGACCGATTAACCGCCCGTGG 2220
Db 2161 TCCATCGCAGAAGCTGCTTTAATAAGGCGAAACAGCGATGACCGATTAACCGCCCGTGG 2220
Qy 2221 GCATGTTCCAAACATCGACACCAAGTGAATTAATGTTGTTACCGGTACTGCGGACCTTC 2280
Db 2221 GCATGTTCCAAACATCGACACCAAGTGAATTAATGTTGTTACCGGTACTGCGGACCTTC 2280
Qy 2281 AAGGGTCAACCATCCAAACCGTTTGGCGTCTGAGCGCAGGTATTAACCGCGCCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTTGGCGTCTGAGCGCAGGTATTAACCGCGCCAGT 2340
Qy 2341 CCCAAACAAAGAGCTGGCAAAAGGTTTCTCGAATACTATCTGCTGATGATGAAGTCTG 2400
Db 2341 CCCAAACAAAGAGCTGGCAAAAGGTTTCTCGAATACTATCTGCTGATGATGAAGTCTG 2400
Qy 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGGGAAAGATCCACGTTATGCGGCCACCATGGAAAAACCGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGGAAAGATCCACGTTATGCGGCCACCATGGAAAAACCGCCAGAAAGGTGAAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGGCTTCTGCTATGCGCTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGGCTTCTGCTATGCGCTGCGTACTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGGTCTGACGACTGTGATGAAGCCCTGAAAGACCGCGACACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTCTGACGACTGTGATGAAGCCCTGAAAGACCGCGACACTAATTCGAGCTCG 2640
Qy 2641 AACACAAACAAATTAACAACTACCAACCAACCTCGGATCGAGGAGGATTTTTCAGAAATTC 2700
Db 2641 AACACAAACAAATTAACAACTACCAACCAACCTCGGATCGAGGAGGATTTTTCAGAAATTC 2700
Qy 2701 GGATCTCT 2709
Db 2701 GGATCTCT 2709

RESULT 9

US-10-343-859-9
; Sequence 9, Application US/10343859
; Publication No. US20040110161A1
; GENERAL INFORMATION:
; APPLICANT: Nanogen Recognomics GMBH
; TITLE OF INVENTION: Method for detecting mutations in
; FILE OF INVENTION: nucleotide sequences
; FILE REFERENCE: 612,406-033
; CURRENT FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: US/10/343,859
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: PCT/EP01/08127
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 10038237.1
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 9191
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: pMALc2x-mutS
; OTHER INFORMATION: plasmid
US-10-343-859-9

Query Match 81.8%; Score 2698; DB 19; Length 9191;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2701; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy	1	CGGACACCATCGAATGGTGCAGAAACCTTTTCGGGGTATGGCATGATGAGCCCGGGAAGAGA	60
Db	1	CGGACACCATCGAATGGTGCAGAAACCTTTTCGGGGTATGGCATGATGAGCCCGGGAAGAGA	60
Qy	61	GTCATTTCAGGGTGGTGAATGTGAACAGTAACTATACGATCTCCAGAGTATGCGG	120
Db	61	GTCAATTTCAGGGTGGTGAATGTGAACAGTAACTATACGATCTCCAGAGTATGCGG	120
Qy	121	GTGCTCTTATCAGACCGTTTCCCGGGTGGTGAACAGGCGACGCACTTCTCGGAAAA	180
Db	121	GTGCTCTTATCAGACCGTTTCCCGGGTGGTGAACAGGCGACGCACTTCTCGGAAAA	180
Qy	181	CGCGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTTACATTTCCCAACCGCGTGGCAC	240
Db	181	CGCGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTTACATTTCCCAACCGCGTGGCAC	240
Qy	241	AACAACTGGCGGGCAACACAGTCTGCTGATGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
Db	241	AACAACTGGCGGGCAACACAGTCTGCTGATGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
Qy	301	ACGCGCGTTCGCAAAATTTGCGGGCGATTTAAATCTCGCGCCGATCAACTGGGTGGCAGCG	360
Db	301	ACGCGCGTTCGCAAAATTTGCGGGCGATTTAAATCTCGCGCCGATCAACTGGGTGGCAGCG	360
Qy	361	TGGTGGTTCGATGTAGAACGAAAGCGGCGTGAAGCCCTGTAAAGCGGCGTGCACAATC	420
Db	361	TGGTGGTTCGATGTAGAACGAAAGCGGCGTGAAGCCCTGTAAAGCGGCGTGCACAATC	420
Qy	421	TTCTCGGCAACCGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACAGGATGCCA	480
Db	421	TTCTCGGCAACCGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACAGGATGCCA	480
Qy	481	TTGCTGTGGAAGCTGCCTCAGTAAATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA	540
Db	481	TTGCTGTGGAAGCTGCCTCAGTAAATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA	540
Qy	541	CACCCATCAACAGTATTTCTCCCATGAAGACGCTACGCGTGGCGGTGGAGCATC	600
Db	541	CACCCATCAACAGTATTTCTCCCATGAAGACGCTACGCGTGGCGGTGGAGCATC	600
Qy	601	TGGTGCATTTGGTCAACGAAATTCGGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG	660
Db			

Db	601	TGGTGCATTTGGTCAACGAAATTCGGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG	660
Qy	661	CGGCTCTGCGTCTGGCTGGTGGCATAAATCTCAGTTCGCAATCAAAATTCAGCCGATAG	720
Db	661	CGGCTCTGCGTCTGGCTGGTGGCATAAATCTCAGTTCGCAATCAAAATTCAGCCGATAG	720
Qy	721	CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCAATGCAAAATGCTGA	780
Db	721	CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCAATGCAAAATGCTGA	780
Qy	781	ATGAGGCGATCGTTCCTCCACTCGCATGCTGGTTGCGCAACGATCGATGGCGCGCAAA	840
Db	781	ATGAGGCGATCGTTCCTCCACTCGCATGCTGGTTGCGCAACGATCGATGGCGCGCAAA	840
Qy	841	TGCGCGCATTTACCGAGTCCGGGTCGGGATGTCGGTATCTCGGTAGTGGGATAGC	900
Db	841	TGCGCGCATTTACCGAGTCCGGGTCGGGATGTCGGTATCTCGGTAGTGGGATAGC	900
Qy	901	ACGATACCGGAAGACAGCTCATGTTATATCCCGCGTAAACCAACCAATCAACAGGATTTTC	960
Db	901	ACGATACCGGAAGACAGCTCATGTTATATCCCGCGTAAACCAACCAATCAACAGGATTTTC	960
Qy	961	GCCTGCTGGGCAAAACCAAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGGCGTGA	1020
Db	961	GCCTGCTGGGCAAAACCAAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGGCGTGA	1020
Qy	1021	AGGGCAATCAGCTGTTGCCCTCTCATCTGGTGAAGAAAAACCAACCTTGGCGCCCAATA	1080
Db	1021	AGGGCAATCAGCTGTTGCCCTCTCATCTGGTGAAGAAAAACCAACCTTGGCGCCCAATA	1080
Qy	1081	CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGCTGCACGACAGGTTT	1140
Db	1081	CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGCTGCACGACAGGTTT	1140
Qy	1141	CCGACTCGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTACTCATTTAG	1200
Db	1141	CCGACTCGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTACTCATTTAG	1200
Qy	1201	GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAAGCGGTGCAACCAATGCTTCGGC	1260
Db	1201	GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCAAGCGGTGCAACCAATGCTTCGGC	1260
Qy	1261	TCAGGCGCATTCGGAAGCTGTGATGCGCTGTGCGAGTCTGTAATCATCTGCATTAATTCG	1320
Db	1261	TCAGGCGCATTCGGAAGCTGTGATGCGCTGTGCGAGTCTGTAATCATCTGCATTAATTCG	1320
Qy	1321	TGTCGCTCAAGCGCACCTCCCGTTCTGGAATAATGTTTTTTCGCGCGACATCATACGGTT	1380
Db	1321	TGTCGCTCAAGCGCACCTCCCGTTCTGGAATAATGTTTTTTCGCGCGACATCATACGGTT	1380
Qy	1381	CTGGCAAAATATTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTTGGA	1440
Db	1381	CTGGCAAAATATTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTTGGA	1440
Qy	1441	ATTGTGAGCGGATTAACAATTTTCAACAGGAAACCGCAGTCCGTTTAGTGTGTTTTCACGA	1500
Db	1441	ATTGTGAGCGGATTAACAATTTTCAACAGGAAACCGCAGTCCGTTTAGTGTGTTTTCACGA	1500
Qy	1501	GCACCTTCAACCAAGGACCATAGATTATGAACTGGAAGAGGTAAATCTGTAATCTGG	1560
Db	1501	GCACCTTCAACCAAGGACCATAGATTATGAACTGGAAGAGGTAAATCTGTAATCTGG	1560
Qy	1561	ATTAACCGCGATTAAGGCTATAACCGCTCTCGCTGAAGTCCGTAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAACCGCGATTAAGGCTATAACCGCTCTCGCTGAAGTCCGTAAGAAATTCGAGAAAGAT	1620
Qy	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCGCCAGGTT	1680
Db	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCGCCAGGTT	1680
Qy	1681	GGGCAATCTGGGATGGCCCTGACATTAATCTTCTGGGCACACCGCTTGTGGTGGCTAC	1740
Db	1681	GGGCAATCTGGGATGGCCCTGACATTAATCTTCTGGGCACACCGCTTGTGGTGGCTAC	1740


```
QY 1741 GCTCAATCTGGCGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db |||||
QY 1741 GCTCAATCTGGCGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db |||||
QY 1801 CGTTTTACTCGGATGCGGTAGCTTTACACGCGAAGCTGATTTGCTTACCCGATCGCTGT 1860
Db |||||
QY 1801 CGTTTTACTCGGATGCGGTAGCTTTACACGCGAAGCTGATTTGCTTACCCGATCGCTGT 1860
Db |||||
QY 1861 GAAAGCGTTATCGCTGATTTATAACAAAGATCTGCTCCGAACTGCTGCGAACCCGCCAAACCTGGGAA 1920
Db |||||
QY 1861 GAAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCGAACCCGCCAAACCTGGGAA 1920
Db |||||
QY 1921 GAGATCCCGCGCTCGATTAAGAACTGAAGCGAAGGTAAAGAGCGCGCTGATTTCAAC 1980
Db |||||
QY 1981 CTGCAAGAACCTGACTTACCTGGCGGCTGATTTGCTGACGGGGTTATGCTTTCAAG 2040
Db |||||
QY 1981 CTGCAAGAACCTGACTTACCTGGCGGCTGATTTGCTGACGGGGTTATGCTTTCAAG 2040
Db |||||
QY 2041 TATGAAAAACGCAAGTACGACATTAAGAACGCTGGCGGTGGATACGCTGGCGGAAAGCG 2100
Db |||||
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAAGAACGCTGGCGGTGGATACGCTGGCGGAAAGCG 2100
Db |||||
QY 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGCTGG 2220
Db |||||
QY 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGCTGG 2220
Db |||||
QY 2221 GCATGGTCCAAATCGACACACGCAAGTGAATTTATGGTGAACGGTACTCCGACCTTC 2280
Db |||||
QY 2281 AAGGTCACACCATCCAAACCGTTGTTGGCGTGTGAGCGAGGTATTAAGCGCGCCAGT 2340
Db |||||
QY 2281 AAGGTCACACCATCCAAACCGTTGTTGGCGTGTGAGCGAGGTATTAAGCGCGCCAGT 2340
Db |||||
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACCTATCTGCTGACTGATGAAGGTCGT 2400
Db |||||
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACCTATCTGCTGACTGATGAAGGTCGT 2400
Db |||||
QY 2401 GAAGCGGTTAATAAGCAAAACCGCTGGTCCGTAGCGCTGAAGTCTTACGAGGAAAG 2460
Db |||||
QY 2461 TTGGCGAAAGATCCACGTTATTCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db |||||
QY 2461 TTGGCGAAAGATCCACGTTATTCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db |||||
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATCGCGTGGCTACTCGCGTGATCAACGCC 2580
Db |||||
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATCGCGTGGCTACTCGCGTGATCAACGCC 2580
Db |||||
QY 2581 GCCAGCGTGTGAGATGAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db |||||
QY 2581 GCCAGCGTGTGAGATGAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db |||||
QY 2641 AACAAACAAACAAATAACAATAACAACCTCGGGATCGAGGAGGATTCAGAAATTC 2700
Db |||||
QY 2701 GGATCC 2706
Db |||||
QY 2701 GGATCC 2706
Db |||||
```

RESULT 10
US-10-263-153-40
; Sequence 40, Application US/10263153

```
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263.153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 7112
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3189)
; OTHER INFORMATION: pMBP-c2X-Toxop30del11 (52-214aa)
; US-10-263-153-40
```

```
Query Match 81.5%; Score 2690.4; DB 18; Length 7112;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGTGAAAACCTTTTCGGGTATGGCATGATAGCGCCCGAAGAGA 60
Db |||||
QY 61 GTCAATTCAGGCTGGTGAATGTGAAACCACTAGTATACGATGTGCGAGATATGCCG 120
Db |||||
QY 61 GTCAATTCAGGCTGGTGAATGTGAAACCACTAGTATACGATGTGCGAGATATGCCG 120
Db |||||
QY 121 GTGCTCTTTATCAGACCGTTTCCCGGTGTGTAACCAAGCCAGCCAGCCAGCTTCTGCCGAAAA 180
Db |||||
QY 121 GTGCTCTTTATCAGACCGTTTCCCGGTGTGTAACCAAGCCAGCCAGCCAGCTTCTGCCGAAAA 180
Db |||||
QY 181 CCGCGGAAAAAGTGGAGACGCGCATGGCGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
Db |||||
QY 181 CCGCGGAAAAAGTGGAGACGCGCATGGCGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
Db |||||
QY 241 AACAACTGGCGGCAAAACAGTCTGCTGATTTGGCGTTCGCCACTCCAGTCTGGCCCTGC 300
Db |||||
QY 241 AACAACTGGCGGCAAAACAGTCTGCTGATTTGGCGTTCGCCACTCCAGTCTGGCCCTGC 300
Db |||||
QY 301 ACSCGCGTCCGCAAAATTTGTCGGGCGAATTAATTTCTCGCGCCGATCAACTGGGTGCCAGCG 360
Db |||||
QY 301 ACSCGCGTCCGCAAAATTTGTCGGGCGAATTAATTTCTCGCGCCGATCAACTGGGTGCCAGCG 360
Db |||||
QY 361 TGGTGTGTGATGTTAGAACGAGCGGTGGAAGCTTGAAGCGGTGTAAGCGGCGGTGCAAAATC 420
Db |||||
QY 361 TGGTGTGTGATGTTAGAACGAGCGGTGGAAGCTTGAAGCGGTGTAAGCGGCGGTGCAAAATC 420
Db |||||
QY 421 TTCTCTCGCAACCGCTCAGTGGGCTGATCAATTAATTTATCGCTGGATGACCAAGATGCCA 480
Db |||||
QY 421 TTCTCTCGCAACCGCTCAGTGGGCTGATCAATTAATTTATCGCTGGATGACCAAGATGCCA 480
Db |||||
QY 481 TTGCTGTGGAAGCTGCTGCACTAATTTCTCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
Db |||||
QY 481 TTGCTGTGGAAGCTGCTGCACTAATTTCTCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
Db |||||
QY 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGAGCGTACGCGCTGGCGGTGGAGCATC 600
Db |||||
QY 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGAGCGTACGCGCTGGCGGTGGAGCATC 600
Db |||||
QY 601 TGGTGTGATTTGGTTCACAGCAAAATCGCGCTGTTAGCGGCGCCCAATTAAGTCTCTGCTCGG 660
Db |||||
QY 601 TGGTGTGATTTGGTTCACAGCAAAATCGCGCTGTTAGCGGCGCCCAATTAAGTCTCTGCTCGG 660
Db |||||
QY 661 CCGCTGTGCGTGTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db |||||
```

Db 661 |||||CGGCTCTGGCTGGCTAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAAACGGGAAGGACATGGAGTGCATGTCCGGTTTTCAACAAACCAATGCAAAATGCTGA 780
Db 721 CGGAAACGGGAAGGACATGGAGTGCATGTCCGGTTTTCAACAAACCAATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATGTTCCCACTGGCATGCTGGTTCGCAACGATCAGATGCGCTGGCGGCA 840
Db 781 ATGAGGGCATGTTCCCACTGGCATGCTGGTTCGCAACGATCAGATGCGCTGGCGGCA 840
Qy 841 TGGCGCCATTACCGAGTCCGGGCTGGCGGTGGCGATATCTCGTGTAGTGGGATACG 900
Db 841 TGGCGCCATTACCGAGTCCGGGCTGGCGGTGGCGATATCTCGTGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTATTATCCGCGGTTAAACCAACCAATCAACACGAGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTATTATCCGCGGTTAAACCAACCAATCAACACGAGATTTTC 960
Qy 961 GCCTGCTGGGCAACCAAGAGTGGACCGGTTGCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACCAAGAGTGGACCGGTTGCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTGCCCGTCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGCCCGTCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Qy 1081 CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTAAATGAGTGGCAGCACAGGTTT 1140
Db 1081 CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTAAATGAGTGGCAGCACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGGCAGTGGAGCGCAACGCAATTAATGTAGTGTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGGAGCGCAACGCAATTAATGTAGTGTAGTCACTCATTTAG 1200
Qy 1201 GCACATTTCTCATGTTTGCAGCTTATCATGACTGCAGGTGCACCAATGCTTCTGGG 1260
Db 1201 GCACATTTCTCATGTTTGCAGCTTATCATGACTGCAGGTGCACCAATGCTTCTGGG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTGTAATCACTGATTAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTGTAATCACTGATTAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGGTTCTGGATAATGTTTTTGGCGGACATCAATACGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGGTTCTGGATAATGTTTTTGGCGGACATCAATACGTT 1380
Qy 1381 CTGGCAATTTCTGAAATGAGCTGTTGACATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAATTTCTGAAATGAGCTGTTGACATTAATCATCGGCTCGTATATGTTGGA 1440
Qy 1441 ATTGTGAGCGGATAACAATTTCAACAGGAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAATTTCAACAGGAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Qy 1501 GCATTTCAACCAAGGACATAGATATGAAAACTGAAGAGGTAACTGGTAATCTGG 1560
Db 1501 GCATTTCAACCAAGGACATAGATATGAAAACTGAAGAGGTAACTGGTAATCTGG 1560
Qy 1561 ATTAACGGGATAAAGGCTATAACCGTCTCGCTGAAGTGGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTAACGGGATAAAGGCTATAACCGTCTCGCTGAAGTGGTAAGAAATTCAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAATCGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAATCGGAAGAGAAATTCACAGGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACAGCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACAGCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCGGCAACAGCGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCGGCAACAGCGTTCCAGGACAAAGCTGTAT 1800

Db 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCGGCAACAGCGTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGATGCGGTACGTTACAAACGCGAAGCTGATGTTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGATGCGGTACGTTACAAACGCGAAGCTGATGTTTACCCGATCGCTGT 1860
Qy 1861 GAAGGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCGCAAAACCTGGGAA 1920
Db 1861 GAAGGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCGCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAACGGAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAACGGAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGCTACTTCACTGCGCGTGTATGCTGACGCGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGCTACTTCACTGCGCGTGTATGCTGACGCGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAATAAACAACACATGATGACAGACCCGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAATAAACAACACATGATGACAGACCCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACGCGCCGTTG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACGCGCCGTTG 2220
Qy 2221 GCATGTTCCACATCGACACAGCAGCAATTAAGTGAATGATGTTGTTAAACGCTGCGACCTTC 2280
Db 2221 GCATGTTCCACATCGACACAGCAGCAATTAAGTGAATGATGTTGTTAAACGCTGCGACCTTC 2280
Qy 2281 AAGGTCACACCATCAAAACCGTTCGTTGGCGTGTGAGCGAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGTCACACCATCAAAACCGTTCGTTGGCGTGTGAGCGAGGTATTAACGCGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGTTAATAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Db 2401 GAAGCGTTAATAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Qy 2461 TTGGGAAAGATCCACGTATTTGCGCCACCATGGAACACGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGGAAAGATCCACGTATTTGCGCCACCATGGAACACGCCAGAAAGGTGAAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Qy 2581 GCGAGCGTGTGAGTGTGATGAGCCCTGAAAGACGCGCAGACTAAATTCGAGTCTG 2640
Db 2581 GCGAGCGTGTGAGTGTGATGAGCCCTGAAAGACGCGCAGACTAAATTCGAGTCTG 2640
Qy 2641 AACACACAAACAAATAACAAACCAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACACACAAACAAATAACAAACCAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700

RESULT 11
US-10-263-153-30
; Sequence 30, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.

/ TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
/ TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
/ FILE REFERENCE: 6984.US.01
/ CURRENT APPLICATION NUMBER: US/10/263,153
/ CURRENT FILING DATE: 2002-10-22
/ NUMBER OF SEQ ID NOS: 74
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 30
/ LENGTH: 7259
/ TYPE: DNA
/ ORGANISM: Toxoplasma gondii
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1528)...(3336)
/ OTHER INFORMATION: pMBP-c2X-Toxop30del14del18 (83-294aa)
US-10-263-153-30

Query Match 81.5%; Score 2690.4; DB 18; Length 7259;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 1 CGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
DB 1 CGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 CTCATTCAGGGTGGTGAATGTGAACCAAGTAAAGTTATACGATGTCGAGAGTATGCCG 120
DB 61 CTCATTCAGGGTGGTGAATGTGAACCAAGTAAAGTTATACGATGTCGAGAGTATGCCG 120
QY 121 GTGTCCTTTATCAGACCGTGTTCGCCGTGGTGAACCGCCGAGCCAGCCAGCTTTCTCGGAAA 180
DB 121 GTGTCCTTTATCAGACCGTGTTCGCCGTGGTGAACCGCCGAGCCAGCCAGCTTTCTCGGAAA 180
QY 181 CGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTAATTCACCAACCGCGTGGCAC 240
DB 181 CGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTAATTCACCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGAACAGTGGTGTGATTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 300
DB 241 AACAACTGGCGGGAACAGTGGTGTGATTTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 300
QY 301 ACGCGCGTGCAGAAATGTGCGCGCGATTAATCTCGCGCGATTAATCTCGCGCGATTAATCT 360
DB 301 ACGCGCGTGCAGAAATGTGCGCGCGATTAATCTCGCGCGATTAATCTCGCGCGATTAATCT 360
QY 361 TGGTGGTGCATGTGAAGCAAGCGCGCTGGAAGCTGTGAAGCGCGCTGGAAGCGCGCTGGA 420
DB 361 TGGTGGTGCATGTGAAGCAAGCGCGCTGGAAGCTGTGAAGCGCGCTGGAAGCGCGCTGGA 420
QY 421 TTCTCGGCAACCGTCAAGTGGGTGATCAATTAATCTCGCGCGATTAATCTCGCGCGATTAAT 480
DB 421 TTCTCGGCAACCGTCAAGTGGGTGATCAATTAATCTCGCGCGATTAATCTCGCGCGATTAAT 480
QY 481 TTGCTGGAAAGCTGCTGCACTAATGTTCGGGGTATTTCTTGATGTCTCTGACGAGA 540
DB 481 TTGCTGGAAAGCTGCTGCACTAATGTTCGGGGTATTTCTTGATGTCTCTGACGAGA 540
QY 541 CACCCATCAACAGTATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGATC 600
DB 541 CACCCATCAACAGTATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGATC 600
QY 601 TGGTGGCATTTGGGTCAACAGCAATCGCGTGTGGGCGCCCATTAAGTTCTGCTCGG 660
DB 601 TGGTGGCATTTGGGTCAACAGCAATCGCGTGTGGGCGCCCATTAAGTTCTGCTCGG 660
QY 661 CGGCTGCGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 720
DB 661 CGGCTGCGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 720
QY 721 CGGAACGGGAAGGCGATGGAGTGGCATGTCCGGTTTTCAACCAATCAATGCTGA 780
DB 721 CGGAACGGGAAGGCGATGGAGTGGCATGTCCGGTTTTCAACCAATCAATGCTGA 780

QY 781 ATGAGGCGCATGTTCCCACTGCGATGCTGGTTCGCAACGATCAGATGGCGCTGGCGCGCA 840
DB 781 ATGAGGCGCATGTTCCCACTGCGATGCTGGTTCGCAACGATCAGATGGCGCTGGCGCGCA 840
QY 841 TGCAGCGCATTAACCGAGTCCGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 900
DB 841 TGCAGCGCATTAACCGAGTCCGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCTTAAACACCATCAACAGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCTTAAACACCATCAACAGATTTTC 960
QY 961 GCCTGCTGGGCGCAACCGAGGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
DB 961 GCCTGCTGGGCGCAACCGAGGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGCGGTGA 1020
QY 1021 AGGCAATACAGTGTGTCCTGCTCTACTGTGTGAAGAAAAACCCACCTCGCGCCCAATA 1080
DB 1021 AGGCAATACAGTGTGTCCTGCTCTACTGTGTGAAGAAAAACCCACCTCGCGCCCAATA 1080
QY 1081 CGCAACCGCTCTCCCGCGCGCTTGGCGATTCATTAATGCACTGGGACGACAGGTTT 1140
DB 1081 CGCAACCGCTCTCCCGCGCGCTTGGCGATTCATTAATGCACTGGGACGACAGGTTT 1140
QY 1141 CCGACTGGAAGCGGCGAGTGAAGCGCAACCGCAATTAATGTAAGTTAGCTCCTCATTTAG 1200
DB 1141 CCGACTGGAAGCGGCGAGTGAAGCGCAACCGCAATTAATGTAAGTTAGCTCCTCATTTAG 1200
QY 1201 GCACAAATTCATGTTTGAAGCTTATATCATGCTGCACTGCAAGTGCACCAATGCTTTCGG 1260
DB 1201 GCACAAATTCATGTTTGAAGCTTATATCATGCTGCACTGCAAGTGCACCAATGCTTTCGG 1260
QY 1261 TCAGGACGCGCATCGGAAGCTGTGATGGCTGCAAGTGCATTAATGCACTGCAATTAATTCG 1320
DB 1261 TCAGGACGCGCATCGGAAGCTGTGATGGCTGCAAGTGCATTAATGCACTGCAATTAATTCG 1320
QY 1321 TGTGCTCAAGCGCGCTCCCGTTCCTGGATTAATGTTTTCGCGCGACATCAATAAGGTT 1380
DB 1321 TGTGCTCAAGCGCGCTCCCGTTCCTGGATTAATGTTTTCGCGCGACATCAATAAGGTT 1380
QY 1381 CTGGCAATATTTCTGAATGAGCTGTGCAATTAATCATCGCTCGTATTAATGTTGTTGA 1440
DB 1381 CTGGCAATATTTCTGAATGAGCTGTGCAATTAATCATCGCTCGTATTAATGTTGTTGA 1440
QY 1441 ATTGTGAGCGGATAACAAATTTACACAGGAAACAGCGAGTCCGTTTGGTGTGTTTCCACGA 1500
DB 1441 ATTGTGAGCGGATAACAAATTTACACAGGAAACAGCGAGTCCGTTTGGTGTGTTTCCACGA 1500
QY 1501 GCATTTCAACCAAGGACCATAGATTAATGAAACTGAAAGAGGTAAACTGGTAAATCTCG 1560
DB 1501 GCATTTCAACCAAGGACCATAGATTAATGAAACTGAAAGAGGTAAACTGGTAAATCTCG 1560
QY 1561 ATTAACGGGATAAAGCGCTTAACCGTCTCGCTGAAAGTGGTAAAGAAATTCGAGAAGAT 1620
DB 1561 ATTAACGGGATAAAGCGCTTAACCGTCTCGCTGAAAGTGGTAAAGAAATTCGAGAAGAT 1620
QY 1621 ACGGAATTAAGTCAACCGTTCGAGTCCGATTAACCTGGAAGAGAAATTCACCAAGTT 1680
DB 1621 ACGGAATTAAGTCAACCGTTCGAGTCCGATTAACCTGGAAGAGAAATTCACCAAGTT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTCTGCGGACACGACCGCTTGGTGGCTAC 1740
DB 1681 GCGGCAACTGGCGATGGCCCTGACATTAATCTCTGCGGACACGACCGCTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCGTGTGGCTGAAATCACCCGCGCAAGCGTTCAGGACAGCTGTAT 1800
DB 1741 GCTCAATCTGGCGTGTGGCTGAAATCACCCGCGCAAGCGTTCAGGACAGCTGTAT 1800
QY 1801 CGGTTTACCTGGGATCGCGTACGTTTACAAACGCGAAGCTGATTGCTTACCAGTCCGCTGT 1860
DB 1801 CGGTTTACCTGGGATCGCGTACGTTTACAAACGCGAAGCTGATTGCTTACCAGTCCGCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTTATTAACAAAGATCTGCTGCCGAACCCGCCAAACCTCGGAA 1920

```
Db 1861 GAAGGTTATCGTGTATTAACAAGATCTGTCGCGAACCCGCAAAACCTTGGAA 1920
Qy 1921 GAGATCCCGCGCTCGATAAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTCGATAAAGAACTGAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTACCTCGGCGCTGATGCTGCTGACGGGGTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTCGGCGCTGATGCTGCTGACGGGGTATGCGTTCAAG 2040
Qy 2041 TATGAAAAACGCAAGTACGACATTAAGAAGCGTGGCGTGGATAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGAAGCGTGGCGTGGATAACGCTGGCGCGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAAGCGCAACATGAATGACAGACCCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAAGCGCAACATGAATGACAGACCCGATTAC 2160
Qy 2161 TCCATCGCAAGCTGCGCTTAAATAAAGCGCAACATGAATGACAGACCCGATTAC 2220
Db 2161 TCCATCGCAAGCTGCGCTTAAATAAAGCGCAACATGAATGACAGACCCGATTAC 2220
Qy 2221 GCATGCTCAACATCGACACAGCAAGTGAATTAATGTTAAACGCTGACCGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACAGCAAGTGAATTAATGTTAAACGCTGACCGACCTTC 2280
Qy 2281 AAGGTCACACCATCAACACCGTTTGGCGTGTGAGCGAGGTATTAACCGCCCGAGT 2340
Db 2281 AAGGTCACACCATCAACACCGTTTGGCGTGTGAGCGAGGTATTAACCGCCCGAGT 2340
Qy 2341 CGAAACAAAGAGCTGGCAAAAGAGTCTCGAAACATATCTGCTGACTGATGAAGGCTG 2400
Db 2341 CGAAACAAAGAGCTGGCAAAAGAGTCTCGAAACATATCTGCTGACTGATGAAGGCTG 2400
Qy 2401 GAAGCGGTTAATAAAGACAAACCGTGGCGTGGAGCGTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCGTGGCGTGGAGCGTGAAGTCTTACGAGGAAG 2460
Qy 2461 TTGGCGAAAGATCCACGATTTGCGCCACCACTGAGAAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTTGCGCCACCACTGAGAAACCGCCAGAAAGGTGAATCATG 2520
Qy 2521 CGAAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTGATGCGGTGATCAACGCC 2580
Db 2521 CGAAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTGATGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACAAACAAACAAATAACAATAACAACCACTCGGGATCGAGGAAGGATTTTCAAGATTC 2700
Db 2641 AACAAACAAACAAATAACAATAACAACCACTCGGGATCGAGGAAGGATTTTCAAGATTC 2700
```

RESULT 12

```
US-10-263-153-35
; Sequence 35, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; FILE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.O1
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 35
; LENGTH: 7322
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528) ... (3399)
; OTHER INFORMATION: pMBP-c2X-Toxop30del10 (52-284aa)
US-10-263-153-35
```

```
Query Match 81.5%; Score 2690.4; DB 18; Length 7322;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
Qy 1 CCGACACCATCGAATGTTGCAAAACCTTTCCGCGTATGGCATGATAGCCCGGGAAGA 60
Db 1 CCGACACCATCGAATGTTGCAAAACCTTTCCGCGTATGGCATGATAGCCCGGGAAGA 60
Qy 61 GTCAATTCAGGGTGGTGAATGTGAACCCAGTAAAGTTATACGATGTCGAGAGTATCGG 120
Db 61 GTCAATTCAGGGTGGTGAATGTGAACCCAGTAAAGTTATACGATGTCGAGAGTATCGG 120
Qy 121 GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACCCAGCCAGCCACCGTTTCTGCAAAA 180
Db 121 GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACCCAGCCAGCCACCGTTTCTGCAAAA 180
Qy 181 CGCGGAAAAAGTGAAGCGCGCATGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGCGCATGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
Qy 241 AACAACTGGCGGGCAACAGTCTGCTGATGTTGGGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGCAACAGTCTGCTGATGTTGGGTTGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGCGCGTCTGCAAAATGTCGCGCGGATTAATCTCGCGCGATCAATCTGGGTGCGAGCG 360
Db 301 ACGCGCGTCTGCAAAATGTCGCGCGGATTAATCTCGCGCGATCAATCTGGGTGCGAGCG 360
Qy 361 TGGTGGTGTGATGAAGCGAGCGCGCTGAAGCGCTGAAGCGCGGTGCAATC 420
Db 361 TGGTGGTGTGATGAAGCGAGCGCGCTGAAGCGCTGAAGCGCGGTGCAATC 420
Qy 421 TTCTCGCGCAACCGGTCAAGTGGGCTGATTAATTAATATATATATATATATATATATAT 480
Db 421 TTCTCGCGCAACCGGTCAAGTGGGCTGATTAATTAATATATATATATATATATATATAT 480
Qy 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGGCTTATTTCTTGATGCTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGGCTTATTTCTTGATGCTCTGACCCAGA 540
Qy 541 CACCCATCAACAGATTAATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGATTAATTTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Qy 601 TGGTGCATTTGGGTACACAGCAAAATCGGCGTGTAGCGGGCCCATTAAGTTCTGTCTCG 660
Db 601 TGGTGCATTTGGGTACACAGCAAAATCGGCGTGTAGCGGGCCCATTAAGTTCTGTCTCG 660
Qy 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGAAACGGGAAGGGAGCTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGAAACGGGAAGGGAGCTGGAGTGCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATCTGTTCCCACTCGCATGCTGGTGTGCAACAGCATAGATGCGGTGGCGCAA 840
Db 781 ATGAGGGCATCTGTTCCCACTCGCATGCTGGTGTGCAACAGCATAGATGCGGTGGCGCAA 840
Qy 841 TGCGGCCATTAACGAGTCCGGGCTGCGGTTGGTGGCGGATATCTCGTAGTGGGATAG 900
Db 841 TGCGGCCATTAACGAGTCCGGGCTGCGGTTGGTGGCGGATATCTCGTAGTGGGATAG 900
```

QY 901 ACGATACCGAAGACAGCTCATGTATATATCCCGCGTTAAACCAACCATCAAAAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTATATATCCCGCGTTAAACCAACCATCAAAAGGATTTTC 960
QY 961 GCCTGCTGGGGCAAAACACAGCTGAGCCGCTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGGCAAAACACAGCTGAGCCGCTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCGTCTACCTGGTCAAAAGAAACCAACCCCTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCCGTCTACCTGGTCAAAAGAAACCAACCCCTGGCGGCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGTGGCAGCACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGTGGCAGCACAGGTTT 1140
QY 1141 CCGGACTGGAAGCGGGGAGTGAAGCGCAACGCAATTAATGTAAGTTAGCTCACTCATTTAG 1200
Db 1141 CCGGACTGGAAGCGGGGAGTGAAGCGCAACGCAATTAATGTAAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAATTCATGTTTGACAGCTTATCATGACTGCACGCTGACCAATGCTTCTGCG 1260
Db 1201 GCACAATTCATGTTTGACAGCTTATCATGACTGCACGCTGACCAATGCTTCTGCG 1260
QY 1261 TCAGGACGCAATCGGAAGCTGTGGTATGCTGTGAGGTGCGTAATCACTGCATAATTCG 1320
Db 1261 TCAGGACGCAATCGGAAGCTGTGGTATGCTGTGAGGTGCGTAATCACTGCATAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACCTCCGCTTCTGGATATGTTTTCGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACCTCCGCTTCTGGATATGTTTTCGCGCGACATCAATACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAAAATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
QY 1441 ATTGTAGCGGATACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTTTTCACGA 1500
Db 1441 ATTGTAGCGGATACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTTTTCACGA 1500
QY 1501 GCACCTTACCAACAGGACCATAGATTATGAACTGAAGAGTAACTGATGTTATCTGG 1560
Db 1501 GCACCTTACCAACAGGACCATAGATTATGAACTGAAGAGTAACTGATGTTATCTGG 1560
QY 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTTCGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTTCGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGATAACTGGAAGAGAAATTCCTCACAGTT 1680
Db 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGATAACTGGAAGAGAAATTCCTCACAGTT 1680
QY 1681 CGCGCAACTGGCGATGGCCCTGACATATCTTCTGGGCACACGACCGCTTGGTGGCTAC 1740
Db 1681 CGCGCAACTGGCGATGGCCCTGACATATCTTCTGGGCACACGACCGCTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGACAAAGGTTCCAGGACAAAGCTGAT 1800
Db 1741 GCTCAATCTGGCCCTGTTGGCTGAAATCACCCCGACAAAGGTTCCAGGACAAAGCTGAT 1800
QY 1801 CCGTTTACCTGGGATGCGTACGTTTACACGGCAAGCTGATGTTTACCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGTACGTTTACACGGCAAGCTGATGTTTACCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCGCCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCGCCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980

QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGANTTCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGANTTCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTAGTACGACATTAAGACAGTGGCGTGGATAACGCTTGGCGCAAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTAGTACGACATTAAGACAGTGGCGTGGATAACGCTTGGCGCAAAAGCG 2100
QY 2101 GGTCTGACCTTCTCCTGGTTGACCTGATTAATAAAACAAACACATGAATCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCCTGGTTGACCTGATTAATAAAACAAACACATGAATCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAAAGCTGCTTTTAAATAAGGCGAAACAGGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAAAGCTGCTTTTAAATAAGGCGAAACAGGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCAAATCGACACACAGCAAAAGTGAATTTATGTTGTAACGGTACTTGCACACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACACAGCAAAAGTGAATTTATGTTGTAACGGTACTTGCACACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTTGGTGGCGTGTGAGCGCAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTTGGTGGCGTGTGAGCGCAGGTATTAACGCGCCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGATTTCTCGAAAACTATCTGCTGATGATGAGGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGATTTCTCGAAAACTATCTGCTGATGATGAGGTCTG 2400
QY 2401 GAAGCGGTTTAAATAAGACAAACCGCTGGTGGCTGAGCGCTGAACTTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTTAAATAAGACAAACCGCTGGTGGCTGAGCGCTGAACTTTACGAGGAAGAG 2460
QY 2461 TTGCGGAAACATCCAGTATTGCGGCACCATCGAAACCGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGCGGAAACATCCAGTATTGCGGCACCATCGAAACCGCCAGAAAGGTGAAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTACTGCGGTGATCAACGCC 2580
QY 2581 GCAGCGGTGCTGAGACTGTCGATGAAGCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCAGCGGTGCTGAGACTGTCGATGAAGCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACCAACAACAATTAACATTAACAATAACAACCTCGGATCGAGGAGGATTTTCAGAAATTC 2700
Db 2641 AACCAACAACAATTAACATTAACAATAACAACCTCGGATCGAGGAGGATTTTCAGAAATTC 2700

RESULT 13

US-10-263-153-25
; Sequence 25, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.O1
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 7352
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3429)

i: OTHER INFORMATION: pMBP-c2X-Toxop30del4C (52-294aa)
US-10-263-123-25

Query Match 81.5%; Score 2690.4; DB 18; Length 7352;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY	1	CGGACCAATCGAATGGTGGCAAAACCTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA	60
DB	1	CGGACCAATCGAATGGTGGCAAAACCTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA	60
QY	61	GTCAATTCAGGGTGGTGAATGTGAACACCAAGTAACTATACGATGTCGACAGTATGCCG	120
DB	61	GTCAATTCAGGGTGGTGAATGTGAACACCAAGTAACTATACGATGTCGACAGTATGCCG	120
QY	121	GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCACGTTTCTGCGAAAA	180
DB	121	GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCACGTTTCTGCGAAAA	180
QY	181	CGCGGAAAGTGAAGCGGGGATGGCGGAGCTGAATTAACATTCCTCAACCGCGTGGCAC	240
DB	181	CGCGGAAAGTGAAGCGGGGATGGCGGAGCTGAATTAACATTCCTCAACCGCGTGGCAC	240
QY	241	AACAACTGGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
DB	241	AACAACTGGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
QY	301	ACGCGCGTTCGCAAAATTTGCGGGGATTAATACTTCGCGCGGATCAACTGGGTGCCAGCG	360
DB	301	ACGCGCGTTCGCAAAATTTGCGGGGATTAATACTTCGCGCGGATCAACTGGGTGCCAGCG	360
QY	361	TGGTGGTTCGATGATAGTGAACGAGCGGCTGAAAGCCCTGTAAAGCGCGGTGCACAAATC	420
DB	361	TGGTGGTTCGATGATAGTGAACGAGCGGCTGAAAGCCCTGTAAAGCGCGGTGCACAAATC	420
QY	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTATCCGCTGATGACCAAGGATGCCA	480
DB	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTATCCGCTGATGACCAAGGATGCCA	480
QY	481	TTGCTGTGGAAAGTCCTGCAATAATGTTTCGGCGTTATTTCTTGATGTTCTGACCCAGA	540
DB	481	TTGCTGTGGAAAGTCCTGCAATAATGTTTCGGCGTTATTTCTTGATGTTCTGACCCAGA	540
QY	541	CACCAATCAACAGTATATTTTCTCCCATGAAGACGGTACGCGATCGGCGTGGAGCATC	600
DB	541	CACCAATCAACAGTATATTTTCTCCCATGAAGACGGTACGCGATCGGCGTGGAGCATC	600
QY	601	TGGTCGCAATGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTCTGTCGCG	660
DB	601	TGGTCGCAATGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTCTGTCGCG	660
QY	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
DB	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAACGGGAAGCGACGTGAGTGCATGTCGCGTTTCAACAAACCAATGCAATGCTGA	780
DB	721	CGGAACGGGAAGCGACGTGAGTGCATGTCGCGTTTCAACAAACCAATGCAATGCTGA	780
QY	781	ATGAGGGCATCGTTCCCACTGGATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	840
DB	781	ATGAGGGCATCGTTCCCACTGGATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	840
QY	841	TGCGGCGCATTAACAGTCCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG	900
DB	841	TGCGGCGCATTAACAGTCCGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG	900
QY	901	ACGATACCGGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAACAGGATTTTC	960
DB	901	ACGATACCGGAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAACAGGATTTTC	960
QY	961	GCCTGCTGGGGCAACACGAGTGGACCGGTTGCTGCAACTCTCTCAGGCGCAGGCGGTGA	1020
DB	961	GCCTGCTGGGGCAACACGAGTGGACCGGTTGCTGCAACTCTCTCAGGCGCAGGCGGTGA	1020

DB	961	GCCTGCTGGGGCAACACGAGTGGACCGGTTGCTGCAACTCTCTCAGGCGCAGGCGGTGA	1020
QY	1021	AGGCAATCAGCTGTGTCCCTCTCACTGTTGAAAGAAAAACCACTTGGCGCCCAATA	1080
DB	1021	AGGCAATCAGCTGTGTCCCTCTCACTGTTGAAAGAAAAACCACTTGGCGCCCAATA	1080
QY	1081	CGCAACCGCTCTCCCGCGGTTGGCCGATTCAATTAATGCAGCTGGCAGCAGGTTT	1140
DB	1081	CGCAACCGCTCTCCCGCGGTTGGCCGATTCAATTAATGCAGCTGGCAGCAGGTTT	1140
QY	1141	CCGACTCGAAAGCGGCGAGTGAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTAG	1200
DB	1141	CCGACTCGAAAGCGGCGAGTGAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTAG	1200
QY	1201	GCACAAATTCATGTTTGACAGCTTATCATCGACGCGTGCACCAATGCTTCTGGCG	1260
DB	1201	GCACAAATTCATGTTTGACAGCTTATCATCGACGCGTGCACCAATGCTTCTGGCG	1260
QY	1261	TCAGCAGCCATCGGAAGCTGTGTATGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG	1320
DB	1261	TCAGCAGCCATCGGAAGCTGTGTATGGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG	1320
QY	1321	TGTGCTCAAGCGCAGCTCCCGTCTTGATTAATGTTTTTGGCCGACATATAACGGTT	1380
DB	1321	TGTGCTCAAGCGCAGCTCCCGTCTTGATTAATGTTTTTGGCCGACATATAACGGTT	1380
QY	1381	CTGCAAAATTTCTGAAATGAGTGTGCAATTAATCATCGGCTCGTAAATCACTGCATATTCG	1440
DB	1381	CTGCAAAATTTCTGAAATGAGTGTGCAATTAATCATCGGCTCGTAAATCACTGCATATTCG	1440
QY	1441	ATTGTGAGCGGATTAACAAATTTTACAGGAAACAGCCAGTCCGTTTGGTGTTCACGA	1500
DB	1441	ATTGTGAGCGGATTAACAAATTTTACAGGAAACAGCCAGTCCGTTTGGTGTTCACGA	1500
QY	1501	GCATTTACCAACAAAGGACCATAGATTATGAACTGAAAGAGGTAATCTGGTAATCTGG	1560
DB	1501	GCATTTACCAACAAAGGACCATAGCATATGAACTGAAAGAGGTAATCTGGTAATCTGG	1560
QY	1561	ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
DB	1561	ATTAACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCCTCAGGTT	1680
DB	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCCTCAGGTT	1680
QY	1681	CGCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACGACCGCTTGGTGGCTAC	1740
DB	1681	CGCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCACACGACCGCTTGGTGGCTAC	1740
QY	1741	GCTCAATTCGCGCTGTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGTAT	1800
DB	1741	GCTCAATTCGCGCTGTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGTAT	1800
QY	1801	CGGTTTACCTGGGATGCGGTACGTTTAAACGGCAAGCTGATGTTTACCGGATCGCTGTT	1860
DB	1801	CGGTTTACCTGGGATGCGGTACGTTTAAACGGCAAGCTGATGTTTACCGGATCGCTGTT	1860
QY	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCGCAACCGCCCAAAACCTGGGAA	1920
DB	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCGCAACCGCCCAAAACCTGGGAA	1920
QY	1921	GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAGGTAAGAGCGCGCTGATGTTCAAC	1980
DB	1921	GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAGGTAAGAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTGCGCGCTGATGTTGCTGACCGGGGGTTATGCGTTCAAG	2040
DB	1981	CTGCAAGAACCGTACTTCACTGCGCGCTGATGTTGCTGACCGGGGGTTATGCGTTCAAG	2040
QY	2041	TATGAAACCGCAAGTACGACATTAAGACGCTGGGCGTGGATTAACGCTGGCGGAAAGCG	2100
DB	2041	TATGAAACCGCAAGTACGACATTAAGACGCTGGGCGTGGATTAACGCTGGCGGAAAGCG	2100

Qy	2101	GGTCTGACCTTCTCTGGTTGACCTGATTAATAAAACAAACACATGAATGAGACACCGGATTAC	2160
Db	2101	GGTCTGACCTTCTCTGGTTGACCTGATTAATAAAACAAACACATGAATGAGACACCGGATTAC	2160
Qy	2161	TCCATCGCAGAAGCTGCCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGGTGG	2220
Db	2161	TCCATCGCAGAAGCTGCCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCCCGGTGG	2220
Qy	2221	GCATGTGTCACATCTGCACACAGCAAAAGTGAATTTATGTGTAAACGGTACTGCCGACCTTC	2280
Db	2221	GCATGTGCCACATCTGCACACAGCAAAAGTGAATTTATGTGTAAACGGTACTGCCGACCTTC	2280
Qy	2281	AAGGGTCAACCATCAAAACCGTTCTGTTGGCGTGTGAGCGCAGGATTAACGCGCGCCAGT	2340
Db	2281	AAGGGTCAACCATCAAAACCGTTCTGTTGGCGTGTGAGCGCAGGATTAACGCGCGCCAGT	2340
Qy	2341	CCGAAACAAAGAGCTGGCAAAAGAGTTCCTCGAATACTATCTCTGACTGATGAAGTCTTG	2400
Db	2341	CCGAAACAAAGAGCTGGCAAAAGAGTTCCTCGAATACTATCTCTGACTGATGAAGTCTTG	2400
Qy	2401	GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG	2460
Db	2401	GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG	2460
Qy	2461	TTGCGGAAAGATCCACGTATTGCGGCCACCATGGAAAAACGCCAGAAAAGTGAATCATG	2520
Db	2461	TTGCGGAAAGATCCACGTATTGCGGCCACCATGGAAAAACGCCAGAAAAGTGAATCATG	2520
Qy	2521	CCGAACATCCCGCAGATGTCGGCTTTCTGTGTATGCGGTGGTACTCGGTGATCAACGCC	2580
Db	2521	CCGAACATCCCGCAGATGTCGGCTTTCTGTGTATGCGGTGGTACTCGGTGATCAACGCC	2580
Qy	2581	GCCAGCGGTGCTGAGACTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
Db	2581	GCCAGCGGTGCTGAGACTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
Qy	2641	AACAAACAAACAATAACAAATACACACACTCGGATCGAGGGAGGATTTTCAGAAATTC	2700
Db	2641	AACAAACAAACAATAACAAATACACACACTCGGATCGAGGGAGGATTTTCAGAAATTC	2700

RESULT 14

```

US-10-263-153-20
/ Sequence 20, Application US/10263153
/ Publication NO. US20040067239A1
/ GENERAL INFORMATION:
/ APPLICANT: Abbott Laboratories
/ APPLICANT: Maine, Gregory I.
/ APPLICANT: Patel, Chandu B.
/ APPLICANT: Ginsburg, Sanford R.
/ APPLICANT: Bliese, Timothy R.
/ TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
/ TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
/ FILE REFERENCE: 6984.US.01
/ CURRENT APPLICATION NUMBER: US/10/263.153
/ CURRENT FILING DATE: 2002-10-22
/ NUMBER OF SEQ ID NOS: 74
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 20
/ LENGTH: 7370
/ TYPE: DNA
/ ORGANISM: Toxoplasma gondii
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1528)...(3447)
/ OTHER INFORMATION: pMBP-c2X-Toxo30del3C (52-300aa)
US-10-263-153-20

```

Query Match 81.5%; Score 2690.4; DB 18; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy	1	CCGACACCAATCGAATGGTGTCAAAACCTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA	60
Db	1	CCGACACCAATCGAATGGTGTCAAAACCTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA	60
Qy	61	GTCAATTCAGGGTGGTAATGTGAACACAGTAACGTTATACGATGTCGCAGAGTATGCCG	120
Db	61	GTCAATTCAGGGTGGTAATGTGAACACAGTAACGTTATACGATGTCGCAGAGTATGCCG	120
Qy	121	GTGTCCTTTATCAGACCCGTTTCCCGCGTGGTGAAACAGGCCAGCCACAGCTTTCTCGGAAAA	180
Db	121	GTGTCCTTTATCAGACCCGTTTCCCGCGTGGTGAAACAGGCCAGCCACAGCTTTCTCGGAAAA	180
Qy	181	CGCGGGAAAAAGTGGAAAGCGCGGATGGCGGAGCTGAATTAATTCCTCCAACCCGGTGGCAAC	240
Db	181	CGCGGGAAAAAGTGGAAAGCGCGGATGGCGGAGCTGAATTAATTCCTCCAACCCGGTGGCAAC	240
Qy	241	AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACCTTCAAGTCTGGGCCCTTGC	300
Db	241	AACAACTGGCGGGCAAAACAGTCGTTGCTGATTTGGCGTTGCCACCTTCAAGTCTGGGCCCTTGC	300
Qy	301	ACGCGCGTTCGCAAAATTTGTCGGCGGATTAATTCGCGCCGATCAATTCGGGTGCCAGCG	360
Db	301	ACGCGCGTTCGCAAAATTTGTCGGCGGATTAATTCGCGCCGATCAATTCGGGTGCCAGCG	360
Qy	361	TGGTGGTTCGATGCTAGAACGAAGCGCGTCGAAGCCTGTAAAGCGGGGTGCACAATC	420
Db	361	TGGTGGTTCGATGCTAGAACGAAGCGCGTCGAAGCCTGTAAAGCGGGGTGCACAATC	420
Qy	421	TTCTCGCGAACCGCTCAGTGGGCTGATCAATTAACATATCCGCTCGATGACCAAGATGCCA	480
Db	421	TTCTCGCGAACCGCTCAGTGGGCTGATCAATTAACATATCCGCTCGATGACCAAGATGCCA	480
Qy	481	TTGCTGTGGAAGCTCCCTGCACTAATGTTCCGGCGTTATTTCTTGAATGCTCTGACCAAGA	540
Db	481	TTGCTGTGGAAGCTCCCTGCACTAATGTTCCGGCGTTATTTCTTGAATGCTCTCTGACCAAGA	540
Qy	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGCACTGGGCGGTGAGCATC	600
Db	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACCGCACTGGGCGGTGAGCATC	600
Qy	601	TGCTCGCATCGGTCACAGCAAAATCGGCTGTTAGCGGGCCCAATTAAAGTCTCTCTCGG	660
Db	601	TGCTCGCATCGGTCACAGCAAAATCGGCTGTTAGCGGGCCCAATTAAAGTCTCTCTCGG	660
Qy	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Db	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Qy	721	CGGAACGGGAAGGCACTGAGTGCATGTCGCGTTCCTCCAAACCATGCAAAATGCTGA	780
Db	721	CGGAACGGGAAGGCACTGAGTGCATGTCGCGTTCCTCCAAACCATGCAAAATGCTGA	780
Qy	781	ATGAGGGCATCGTTCCCACTCGCATGCTGGTTCGCAACGATCAGATGGCGCTGGCGCAA	840
Db	781	ATGAGGGCATCGTTCCCACTCGCATGCTGGTTCGCAACGATCAGATGGCGCTGGCGCAA	840
Qy	841	TGCGCGCATTCAGAGTCCGGGCTCGCGTTCGCGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGCGCATTCAGAGTCCGGGCTCGCGTTCGCGATATCTCGGTAGTGGGATACG	900
Qy	901	ACGNATCCGAAGACAGCTCAATGTTATATCCGCGGTTAACCAACATCAACAGATTTTC	960
Db	901	ACGNATCCGAAGACAGCTCAATGTTATATCCGCGGTTAACCAACATCAACAGATTTTC	960
Qy	961	GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAAACTCTCTCAGGGCCAGCGGTGA	1020
Db	961	GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAAACTCTCTCAGGGCCAGCGGTGA	1020
Qy	1021	AGGCGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAAAACCACTCTGCGGCCCAATA	1080
Db	1021	AGGCGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAAAACCACTCTGCGGCCCAATA	1080
Qy	1081	CGCAAAACCGCCTCTCCCGCGCGTGGCGGATTCATTAATGAGCTGGCAGCAGCATGTTT	1140

Db 1081 ||||| CGCAAAACGCTCTCCCGCGGCTTGGCGGATTCAATTAATGCAAGTGGACGACAGGTTT 1140
Qy 1141 CCGGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCGGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Qy 1201 GCACAAATTTCTCATGTTTGACAGCTTATCATGCACTGCAAGGTCGACCAATGCTTCTGGG 1260
Db 1201 GCACAAATTTCTCATGTTTGACAGCTTATCATGCACTGCAAGGTCGACCAATGCTTCTGGG 1260
Qy 1261 TCAGGACGCAATCGGAAGCTGTGATGCTGTGAGCTGTAATCACTGCAATCAATTAATTCG 1320
Db 1261 TCAGGACGCAATCGGAAGCTGTGATGCTGTGAGCTGTAATCACTGCAATCAATTAATTCG 1320
Qy 1321 TGTGCTCAAGCGGCACTCCCGTTCTGATTAATGTTTTTGGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCTGATTAATGTTTTTGGCGCGACATCAATACGGTT 1380
Qy 1381 CTGGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATGATGTTGGA 1440
Db 1381 CTGGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATGATGTTGGA 1440
Qy 1441 ATTGTGAGGGATAACAATTTTACACAGAAACAGCCAGTCCGTTTAGTGTTTTTCAGGA 1500
Db 1441 ATTGTGAGGGATAACAATTTTACACAGAAACAGCCAGTCCGTTTAGTGTTTTTCAGGA 1500
Qy 1501 GCATTTCAACCAAGGACCAATTTATGAAAACCTGAAGAGGTAATGTTGTTTCTGG 1560
Db 1501 GCATTTCAACCAAGGACCAATTTATGAAAACCTGAAGAGGTAATGTTGTTTCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGTGAAGTCCGTTAGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGTGAAGTCCGTTAGGTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACGTTGAGCATCCGATTAACCTGGAAGAGAAATTCOCACAGTT 1680
Db 1621 ACCGGAATTAAGTCAACGTTGAGCATCCGATTAACCTGGAAGAGAAATTCOCACAGTT 1680
Qy 1681 GCGCAACTGCGGATGCGCTGACATTAATCTTTGCGGCACAGCCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGCGGATGCGCTGACATTAATCTTTGCGGCACAGCCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTTGGCTGAATATCAACCGGACAAAGCGTTCCAGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTTGGCTGAATATCAACCGGACAAAGCGTTCCAGACAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGATGCGGTAGTTTACACGCGAAGCTGATTTGCTTACCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGGTAGTTTACACGCGAAGCTGATTTGCTTACCGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCGGAAACCCGCCAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCGGAAACCCGCCAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCTCGGATAAGAACTGAAGCGAAAGTGAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCTCGGATAAGAACTGAAGCGAAAGTGAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGAGCGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGAGCGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAACGGCAAGTACGACATTAAGACGTTGGGCTGGATAACCGTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGTTGGGCTGGATAACCGTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACCAATGAATGACAGACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACCAATGAATGACAGACCGGATTAC 2160
Qy 2161 TCCATCGCAAGAGCTGCTTTAATTAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220

Db 2161 TCCATCGCAAGAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCCAAATCGACACAGCAAAAGTGAATTAATGGTTAAACGGTACTCGCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACAGCAAAAGTGAATTAATGGTTAAACGGTACTCGCGACCTTC 2280
Qy 2281 AAGGGTCAACCATCAAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAACCGCCGCAAT 2340
Db 2281 AAGGGTCAACCATCAAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAACCGCCGCAAT 2340
Qy 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTTCTG 2400
Qy 2401 GAAGCGGTTAATAAAGCAAAACCGCTGGGTGCGGTAGCGCTGAAGCTTTTACGAGAAAG 2460
Db 2401 GAAGCGGTTAATAAAGCAAAACCGCTGGGTGCGGTAGCGCTGAAGCTTTTACGAGAAAG 2460
Qy 2461 TTGGCGAAAGATCCACGTTATTGCGCCCACTATGGAACACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATTGCGCCCACTATGGAACACGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGGTGCTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGGTGCTGCTGCTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Qy 2641 AACACACACAAATTAACAATAACAACCAACCTCGGATCGAGGAGGATTTTCAGAAATTC 2700
Db 2641 AACACACACAAATTAACAATAACAACCAACCTCGGATCGAGGAGGATTTTCAGAAATTC 2700

RESULT 15
US-10-263-153-61
; Sequence 61, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984 US 01
; CURRENT APPLICATION NUMBER: US/10/263,153
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 7370
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3447)
; OTHER INFORMATION: pMBP-c2X-ToxoP30MIX1
US-10-263-153-61

Query Match 81.5%; Score 2690.4; DB 18; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGGGTATGGCATGTAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGGGTATGGCATGTAGCGCCCGGAAGAGA 60
Qy 61 GTCATTTACAGGTTGTTGAATGTAACCAAGTAAAGTTATACGATGTCGACAGGATATCGG 120
Db 61 GTCATTTACAGGTTGTTGAATGTAACCAAGTAAAGTTATACGATGTCGACAGGATATCGG 120

121 GTGTCTCTTATCAGACCGTTTTCCGGGTGGTGAACAGGCGCAGCGTTCCTGCGAAA 180
121 GTGTCTCTTATCAGACCGTTTTCCGGGTGGTGAACAGGCGCAGCGTTCCTGCGAAA 180
181 CGCGGGAAGTGGAAAGCGCGGATGGCGGAGCTGAATACATTCCTCAACCGCGTGGCAC 240
181 CGCGGGAAGTGGAAAGCGCGGATGGCGGAGCTGAATACATTCCTCAACCGCGTGGCAC 240
241 AACAACTGGCGGCAAAACAGTCTGCTGATTTGGCGTTGGCCACTCCAGTCTGGCCCTGC 300
241 AACAACTGGCGGCAAAACAGTCTGCTGATTTGGCGTTGGCCACTCCAGTCTGGCCCTGC 300
301 ACSCGCGTGCAGAAATGTGGCGGATTAATCTCGCGCGGATCAACTGGGTGGCAGCG 360
301 ACSCGCGTGCAGAAATGTGGCGGATTAATCTCGCGCGGATCAACTGGGTGGCAGCG 360
361 TGGTGTGTGATGGTGAACAGGCGGCTGAAAGCTGTAAAGCGCGGTGCACAATC 420
361 TGGTGTGTGATGGTGAACAGGCGGCTGAAAGCTGTAAAGCGCGGTGCACAATC 420
421 TTCTCGCGCAACCGGTCACTGGGCTGATCAATTAATCTATCCGCTGGATGACAGGATGCCA 480
421 TTCTCGCGCAACCGGTCACTGGGCTGATCAATTAATCTATCCGCTGGATGACAGGATGCCA 480
481 TTGCTGTGAAGTGCCTGCACTAATGTTCGGCGTTAATTTCTTGATGTCTCTGACCCAGA 540
481 TTGCTGTGAAGTGCCTGCACTAATGTTCGGCGTTAATTTCTTGATGTCTCTGACCCAGA 540
541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGGACTGGCGGTGGAGCATC 600
541 CACCCATCAACAGTATTAATTTCTCCCATGAAGCGGTACGGACTGGCGGTGGAGCATC 600
601 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
601 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
661 CGCGTCTGGTCTGGCTGGGCAATATCTCACTCGCAATCAAAATCAGCGGATAG 720
661 CGCGTCTGGTCTGGCTGGGCAATATCTCACTCGCAATCAAAATCAGCGGATAG 720
721 CGGAAACGGGAAGGACATGGAGTGCATGTCCGGTTTTCAACAAACCATGCAATGCTGA 780
721 CGGAAACGGGAAGGACATGGAGTGCATGTCCGGTTTTCAACAAACCATGCAATGCTGA 780
781 ATGAGGCGCATGTTCCCACTCGGATGCTGTGGTGGCAACGATCAGATGGCGTGGCGCAA 840
781 ATGAGGCGCATGTTCCCACTCGGATGCTGTGGTGGCAACGATCAGATGGCGTGGCGCAA 840
841 TGGCGGCAATACCGAGTCCGGGCTGGCGTGGTGGGATATCTCGGTAGTGGGATACG 900
841 TGGCGGCAATACCGAGTCCGGGCTGGCGTGGTGGGATATCTCGGTAGTGGGATACG 900
901 ACGATACGGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACATCAACAGGATTTTC 960
901 ACGATACGGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACATCAACAGGATTTTC 960
961 GCCTGTGGGCAACAGCGTGAACGCTTGTGCAACTCTCTCAGGCGCAGCGGTGA 1020
961 GCCTGTGGGCAACAGCGTGAACGCTTGTGCAACTCTCTCAGGCGCAGCGGTGA 1020
1021 AGGCAATACAGCTGTTCCCGTCTCACTGTGAAAGAAAAACACCCCTGGCGCCCAATA 1080
1021 AGGCAATACAGCTGTTCCCGTCTCACTGTGAAAGAAAAACACCCCTGGCGCCCAATA 1080
1081 CGCAACCGGCTCTCCCGCGGTGGCCGATTAATGCAAGTGGCAAGCAGGTTT 1140
1081 CGCAACCGGCTCTCCCGCGGTGGCCGATTAATGCAAGTGGCAAGCAGGTTT 1140
1141 CCCGACTGGAAGCGGCGAGTGAACGCAATTAATGTAAGTTAGCTCACTATTAG 1200
1141 CCCGACTGGAAGCGGCGAGTGAACGCAATTAATGTAAGTTAGCTCACTATTAG 1200

1201 GCACAAATCTCATGTTTGCACAGCTTATCATCGACTGCAACGCTGCAACCAATCTTCTGGCG 1260
1201 GCACAAATCTCATGTTTGCACAGCTTATCATCGACTGCAACGCTGCAACCAATCTTCTGGCG 1260
1261 TCAGGACGCCATCCGAAAGCTGTGGTATGGCTGTGAGGTCTGTAATCATCTGCAATAATTCG 1320
1261 TCAGGACGCCATCCGAAAGCTGTGGTATGGCTGTGAGGTCTGTAATCATCTGCAATAATTCG 1320
1321 TGTGCTCAAGCGGCACCTCCCGTTCTGGATAATGTTTTTTTGGCGCGGACATCATACCGGTT 1380
1321 TGTGCTCAAGCGGCACCTCCCGTTCTGGATAATGTTTTTTTGGCGCGGACATCATACCGGTT 1380
1381 CTGCAATAATCTGAAATGAGCTGTGGACAAATTAATCATCTGGCTCGTATAATGTGGGA 1440
1381 CTGCAATAATCTGAAATGAGCTGTGGACAAATTAATCATCTGGCTCGTATAATGTGGGA 1440
1441 ATTCTGAGCGGATAAACAAATTCACACAGGAAACAGCCAGTCCGTTTAAAGTGTGTTTACGA 1500
1441 ATTCTGAGCGGATAAACAAATTCACACAGGAAACAGCCAGTCCGTTTAAAGTGTGTTTACGA 1500
1501 GCATTCACCAACAGGACCATAGATTAATGAAATCTGAAGAGTAAATCTGGTAATCTGG 1560
1501 GCATTCACCAACAGGACCATAGATTAATGAAATCTGAAGAGTAAATCTGGTAATCTGG 1560
1561 ATTAACCGCGATAAAGCTATACCGTCTGCTGAAAGTCCGTTAAGAAATTCGAGAAAGAT 1620
1561 ATTAACCGCGATAAAGCTATACCGTCTGCTGAAAGTCCGTTAAGAAATTCGAGAAAGAT 1620
1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT 1680
1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGGTT 1680
1681 GCGGCACTGGGCAATGGCCCTGACATTAATCTCTGGGCAACAGCCGTTTGGTGGCTAC 1740
1681 GCGGCACTGGGCAATGGCCCTGACATTAATCTCTGGGCAACAGCCGTTTGGTGGCTAC 1740
1741 GCTCAATCTGGCTGTGGCTGAAATCACCCGGAACAAGCGTTCCAGAGCAAGCTGTAT 1800
1741 GCTCAATCTGGCTGTGGCTGAAATCACCCGGAACAAGCGTTCCAGAGCAAGCTGTAT 1800
1801 CGGTTTACTGGGATGCGGTAAGTTTCAACCGCAAGCTGATTTACCCGATCGCTGTT 1860
1801 CGGTTTACTGGGATGCGGTAAGTTTCAACCGCAAGCTGATTTACCCGATCGCTGTT 1860
1861 GAAGCGTATCGCTGATTTATCAACAGATCTGCTGCCGAACCCGCAAAAAACCTGGGAA 1920
1861 GAAGCGTATCGCTGATTTATCAACAGATCTGCTGCCGAACCCGCAAAAAACCTGGGAA 1920
1921 GAGATCCCGCGCTGGATAAGAACTGAAAGCAAGGTAAAGCGCGCTGATGTTCAAC 1980
1921 GAGATCCCGCGCTGGATAAGAACTGAAAGCAAGGTAAAGCGCGCTGATGTTCAAC 1980
1981 CTGCAAGAACCGTACTTCACTGGCGGTGATTTGCTGTGACCGGGGTTATGCGTTCAAG 2040
1981 CTGCAAGAACCGTACTTCACTGGCGGTGATTTGCTGTGACCGGGGTTATGCGTTCAAG 2040
2041 TATGAAAAACGCAAGTACGACATTAAGACGTTGGGCGTGAATACCGCTGGCGCGAAAGCG 2100
2041 TATGAAAAACGCAAGTACGACATTAAGACGTTGGGCGTGAATACCGCTGGCGCGAAAGCG 2100
2101 GGTCTGAACCTTCTGGTTGACCTGATTAATAAACAACATGATGACAGACCGATTTAC 2160
2101 GGTCTGAACCTTCTGGTTGACCTGATTAATAAACAACATGATGACAGACCGATTTAC 2160
2161 TCCATCGGAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
2161 TCCATCGGAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
2221 GCATGGTCCAACATCGACACCAAGCAAGTGAATTAATGTTGTAACGGTACTGCCGACCTTC 2280
2221 GCATGGTCCAACATCGACACCAAGCAAGTGAATTAATGTTGTAACGGTACTGCCGACCTTC 2280
2281 AAGGCTCAACCATCCAAACCGGTTGTTGGCGTGTGAGCGCAGGTATTAAACGGCCGAGT 2340

[illegible]

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 6743.2 Seconds
(without alignments)
18627.964 Million cell updates/sec

Title: US-09-765-555B-18
Perfect score: 3300
Sequence: 1 ccgcacacatcgatgtgc.....acgacgttcggactaacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_ges1:*
9: gb_ges2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	778	23.6	794	5	BQ751936
2	770.6	23.4	779	7	CK394523
3	768	23.3	769	7	CK118014
4	762	23.1	793	5	BQ751655
5	755.4	22.9	806	5	BQ751220
6	684.8	20.8	689	6	CB863541
7	676.4	20.5	712	1	AL038548
8	676	20.5	710	8	AY080106
9	647.4	19.6	660	1	AL044483
10	641	19.4	641	4	BJ063872
11	615.8	18.7	624	6	CB863814
12	612	18.5	794	1	AL045353
13	603	18.3	752	9	CL655575
14	589.8	17.9	608	8	AY080094
15	586.4	17.8	589	1	AL037742
16	581.2	17.6	501	1	AV594538
17	579	17.5	579	1	AL043868
18	570	17.3	782	7	CK394246
19	569	17.2	774	1	AL045337
20	563.2	17.1	585	9	CL658017
21	562.6	17.0	723	1	AL039416
22	549.4	16.6	753	1	AL045341
23	546.8	16.6	782	1	AL037051
24	545.4	16.5	547	6	CA895956

C 25	545	16.5	648	7	CK394349	hggada2D0
C 26	540.2	16.4	569	1	AL039077	DKF2p566G
C 27	534.4	16.2	644	6	CB862524	HH04B08u
C 28	534	16.2	734	1	AL048427	DKF2p588H
C 29	529	16.0	740	1	AL042909	DKF2p434J
C 30	528	16.0	720	6	CA881974	CA881974
C 31	523.2	15.9	528	1	AL039811	K0994C10-
C 32	519.4	15.7	521	6	CA890032	DKF2p566O
C 33	519	15.7	607	6	CB862099	B0156F05-
C 34	516.4	15.6	518	6	CA886082	HH06A02Y
C 35	510.2	15.5	686	1	AL044407	B0124F04-
C 36	509.4	15.4	511	6	CA895436	DKF2p434D
C 37	498	15.1	498	1	AL039076	B0192A07-
C 38	494.8	15.0	1025	1	AL038025	CA895436
C 39	487.6	14.8	498	7	CK394397	DKF2p566G
C 40	477.6	14.5	491	6	CB862715	hggada2H1
C 41	465.8	14.1	679	1	AL039128	HH03F14u
C 42	462.8	14.0	490	1	AL039589	DKF2p566K
C 43	458	13.9	458	6	CA887583	DKF2p434D
C 44	456.4	13.8	471	1	AL039649	B0137H04-
C 45	449.2	13.6	612	6	CB863409	CA887583

ALIGNMENTS

RESULT 1
BQ751936
LOCUS EST632499 DSCT Colletotrichum trifolii cDNA clone pDSCT8-67, mRNA 794 bp linear EST 18-JUL-2002
DEFINITION sequence.
ACCESSION BQ751936
VERSION BQ751936.1 GI:21907341
KEYWORDS EST.
SOURCE Colletotrichum trifolii
ORGANISM Colletotrichum trifolii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
mitosporic Phyllachoraceae; Colletotrichum.
REFERENCE 1. (bases 1 to 794)
Samad D.A., Dickman, M., Town, C.D., Van Aken, S., Utterback, T.,
Cheung, F. and Fraser, C.M.
ESTs from mycelia of Colletotrichum trifolii race 1
Other ESTs: EST632498
Contact: Deborah A. Samad
Department of Plant Pathology
University of Minnesota
495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
Tel: 612 625 1243
Fax: 651 649 5058
Email: debby@puccini.crl.umn.edu
TIGR sequence name: MTSAH67TV More information is available at:
www.medicago.org
Seq primer: 19cA AtA CgA Ctc Act AtA 999 C).

FEATURES
source
1..794
/organism="Colletotrichum trifolii"
/mol_type="mRNA"
/strain="race 1"
/db_xref="taxon:5466"
/clone="pDSCT8-67"
/tissue_type="mycelia"
/dev_stage="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
/lab_host="DH5alpha"
/clone_lib="DSCt"
/note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; Isolate: 2ap2 ; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gtl1 from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce.

E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN

Query Match	23.6%;	Score 778;	DB 5;	Length 794;	
Best Local Similarity	98.7%;	Pred. No. 3.4e-219;			
Matches 784;	Conservative 0;	Mismatches 10;	Indels 0;	Gaps 0;	
QY	3	GACACATCGATGTCGCAAAACCTTTCGCGTATGGCATGATAGCGCCGGAAGAGT	62		
Db	1	GACATTTTGAATGCGCGCAAAACCTTTCGCGTATGGCATGATAGCGCCGGAAGAGT	60		
QY	63	CAATTTCAGGGTGGTGAATGTGAAACAGTAAAGTATACGATGTCGAGATGATCGCGGT	122		
Db	61	CAATTTCAGGGTGGTGAATGTGAAACAGTAAAGTATACGATGTCGAGATGATCGCGGT	120		
QY	123	GTCTCTTATCAGACCGTTTCCGCGTGTGAAACAGGCGCAGCGTTCCTGCGAAACG	182		
Db	121	GTCTCTTATCAGACCGTTTCCGCGTGTGAAACAGGCGCAGCGTTCCTGCGAAACG	180		
QY	183	CGGGAAGAGTGGAGCGCGATGCGGAGCTGAATTAATTCACACCGCGTGACAA	242		
Db	181	CGGGAAGAGTGGAGCGCGATGCGGAGCTGAATTAATTCACACCGCGTGACAA	240		
QY	243	CAACTGGCGGCAACAGCTGTTGCTGATGCGGTGCGCACTCCAGTCTGCGCTGCAC	302		
Db	241	CAACTGGCGGCAACAGCTGTTGCTGATGCGGTGCGCACTCCAGTCTGCGCTGCAC	300		
QY	303	CGCGCTGCGCAAAATGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCGTG	362		
Db	301	CGCGCTGCGCAAAATGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCGTG	360		
QY	363	GTGGTGTGATGATGAGACGAGCGGTGCAAGCTGTAAGCGCGGTGCAATCTT	422		
Db	361	GTGGTGTGATGATGAGACGAGCGGTGCAAGCTGTAAGCGCGGTGCAATCTT	420		
QY	423	CTCGCGCAACCGTCACTGGGCTGATCAATTAATCTCGCTGGATGACCGAGTGCAT	482		
Db	421	CTCGCGCAACCGTCACTGGGCTGATCAATTAATCTCGCTGGATGACCGAGTGCAT	480		
QY	483	GCTGTGAGTGGCTGACATAATGTTCCGCGTATTTCTTGATGTCCTGACAGACA	542		
Db	481	GCTGTGAGTGGCTGACATAATGTTCCGCGTATTTCTTGATGTCCTGACAGACA	540		
QY	543	CCCATCAACAGTATTATTTCTCCATGAACGCTAGCGACTGGGGTGGACATCTG	602		
Db	541	CCCATCAACAGTATTATTTCTCCATGAACGCTAGCGACTGGGGTGGAGCATCTG	600		
QY	603	GTGCAATTTGGTTCACCGCAAAATCGCGTGTAGCGGCGCCATTAAGTTCCTCGCGG	662		
Db	601	GTGCAATTTGGTTCACCGCAAAATCGCGTGTAGCGGCGCCATTAAGTTCCTCGCGG	660		
QY	663	CGTCTGGTCTGGCTGGTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAGC	722		
Db	661	CGTCTGGTCTGGCTGGTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAGC	720		
QY	723	GACCGGAAGCGACTGGAGTGCCATGTCGCGTTCACCAACCAATGCAATGCTGAAT	782		
Db	721	GACCGGAAGCGACTGGAGTGCCATGTCGCGTTCACCAACCAATGCAATGCTGAAT	780		
QY	783	GAGGGCATCGTTCC	796		
Db	781	GAGGGCATCGTTCC	794		

RESULT 2
CK394523/c
LOCUS
DEFINITION hggada4D08 Gland Cell Amplified cDNA Library Heterodera glycines 779 bp mRNA linear EST 30-DEC-2003

cdna, mRNA sequence.
CK394523
CK394523.1 GI:40389794
EST.
Heterodera glycines
Heterodera glycines
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
Tylenchoidea; Heteroderidae; Heteroderinae; Heterodera.
1 (bases 1 to 779)
Wang,X., Allen,R., Gao,B., Goellner,M., Maier,T., Baum,T.,
Hussey,R. and Davis,E.
Submission: Wang, Allen, Gao, Goellner, Maier, Baum, Hussey, Davis
Unpublished (2003)
Contact: Tom Maier
Department of Plant Pathology, Baum Lab
Iowa State University
351 Bessey Hall, Ames, IA 50011, USA
Tel: 515-294-8854
Fax: 515-294-9420
Email: tmaier@iastate.edu
Heterodera glycines Gland Cell Amplified cDNA Library, single pass
sequence.
Location/Qualifiers
1..779
/organism="Heterodera glycines"
/mol_type="mRNA"
/db_xref="taxon:51029"
/cell_type="gland"
/dev_stage="mixed parasitic juvenile"
/clone_lib="Gland Cell Amplified cDNA Library"
/note="Organ: gland cell; Vector: pSportII"

ORIGIN

Query Match	23.4%;	Score 770.6;	DB 7;	Length 779;	
Best Local Similarity	99.4%;	Pred. No. 5.3e-217;			
Matches 773;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;	
QY	87	CCAGTAACGTTATACGATGTCGAGAGTATGCCGGTGTCTCTTATCAGACGCTTCCCGC	146		
Db	779	CCAGTAACGTTATACGATGTCGAGAGTATGCCGGTGTCTCTTATCAGAGTTTCNCCG	720		
QY	147	GTGGTGAACAGCGCAGCCACGTTTCTCGAAACCGCGGAAAAAGTGGAAAGCGGCGATG	206		
Db	719	GTGGTGAACAGCGCAGCCACGTTTCTCGAAACCGCGGAAAAAGTGGAAAGCGGCGATG	660		
QY	207	GCGGAGCTGAATTACATTTCCAAACCGCTGGCACAACTCTGGGGGCAACAGTCGTTG	266		
Db	659	GCGGAGCTGAATTACATTTCCAAACCGCTGGCACAACTCTGGGGGCAACAGTCGTTG	600		
QY	267	CTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCACGCGCGTCCGAAATTTGTCGGCGG	326		
Db	599	CTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGCACGCGCGTCCGAAATTTGTCGGCGG	540		
QY	327	ATTAATCTCGCGCCCATCACTGGGTGTCAGCGTGGTGTGATGTGATGAGCAAGC	386		
Db	539	ATTAATCTCGCGCCCATCACTGGGTGTCAGCGTGGTGTGATGTGATGAGCAAGC	480		
QY	387	GCGGTGCAAGCCCTGTAAGCGCGTGCAATCTTCTCGCGCAACGCTCAGTCAGTGGCTG	446		
Db	479	GCGGTGCAAGCCCTGTAAGCGCGTGCAATCTTCTCGCGCAACGCTCAGTCAGTGGCTG	420		
QY	447	ATCATTAACTATCCGCTGGATGACCAAGATGCCATTGCTGTGGAAGCTGCTGCATTAAT	506		
Db	419	ATCATTAACTATCCGCTGGATGACCAAGATGCCATTGCTGTGGAAGCTGCTGCATTAAT	360		
QY	507	GTTCGGCGTTATTTCTTGATGTCTTGACGACAGACCATCAACAGTATATTTTCTCC	566		
Db	359	GTTCGGCGTTATTTCTTGATGTCTTGACGACAGACCATCAACAGTATATTTTCTCC	300		
QY	567	CATGAAGACGGTACGCGACTGGGCGTGGAGCATCTGGTCGATTGGGTCAACGAGCAATC	626		
Db	299	CATGAAGACGGTACGCGACTGGGCGTGGAGCATCTGGTCGATTGGGTCAACGAGCAATC	240		

COMMENT

Other ESTs: EST632217
 Contact: Deborah A. Samac
 Department of Plant Pathology
 University of Minnesota
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
 Tel: 612 625 1243
 Fax: 651 649 5058
 Email: debbys@pucini.crl.umn.edu
 TIGR sequence name: MTSAF59TV More information is available at:
 www.medicago.org
 Seq primer: (gtA AtA CgA Ctc Act AtA 99g C).
 Location/Qualifiers

FEATURES

source

1. .793
 /organism="Colletotrichum trifolii"
 /mol_type="mRNA"
 /strain="race 1"
 /db_xref="taxon:5466"
 /clone="pDST6-59"
 /tissue_type="mycelia"
 /dev_stage="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
 /lab_host="DH5alpha"
 /clone_lib="DST"
 /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2sp2 ; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gt11 from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN

Query Match 23.1%; Score 762; DB 5; Length 793;
 Best Local Similarity 99.1%; Pred. No. 1.9e-214;
 Matches 777; Conservative 0; Mismatches 5; Indels 2; Gaps 1;
 3 GACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATATAGCGCCGGAAGAGAGT 62
 1 GACATCTTCGAATGGCGCAAAACCTTTCGCGGTATGGCATATAGCGCCGGAAGAGAGT 60
 63 CAATTCAGGGTGGTGAATGTGAACCCAGTAACGTTATACGATGCGCAGAGTATGCCGGT 122
 61 CAATTCAGGGTGGTGAATGTGAACCCAGTAACGTTATACGATGCGCAGAGTATGCCGGT 120
 123 GTCTCTTATCAGACCGTTTCCGCGGTGTGAACCCAGTACGTTATACGATGCGCAGAGTATGCCGGT 182
 121 GTCTCTTATCAGACCGTTTCCGCGGTGTGAACCCAGTACGTTATACGATGCGCAGAGTATGCCGGT 180
 183 CGGGAAAAAGTGGAAAGCGGGGATCGCGAGTGAATTAACATTCCTCAACCGCGTGGGACAA 242
 181 CGGGAAAAAGTGGAAAGCGGGGATCGCGAGTGAATTAACATTCCTCAACCGCGTGGGACAA 240
 243 CAATCGCGGCAACAGTGTGCTGATGGCGGTGCGACCTCCAGTTCGGCCCTGCGAC 302
 241 CAATCGCGGCAACAGTGTGCTGATGGCGGTGCGACCTCCAGTTCGGCCCTGCGAC 300
 303 GCGCGGTGCGAAATGTTCGCGGGGATTAATATTCGCGCGGATCAACTGGGTGGCGCGGTG 362
 301 GCGCGGTGCGAAATGTTCGCGGGGATTAATATTCGCGCGGATCAACTGGGTGGCGCGGTG 360
 363 GTGGTGTGATGTGAGAACAGCGGGGTGCAAGCGCTGTAAGCGGGGTGCAATCTT 422
 361 GTGGTGTGATGTGAGAACAGCGGGGTGCAAGCGCTGTAAGCGGGGTGCAATCTT 420
 423 CTCGCGCAACCGGTGAGTGGGCTGATCAATTAATATTCGCGGTGATGATCAGGAGTGCATT 482
 421 CTCGCGCAACCGGTGAGTGGGCTGATCAATTAATATTCGCGGTGATGATCAGGAGTGCATT 480

QY 483 GCTGTGGAAGCTGCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTTCTTGACACGACA 542
 DB 481 GCTGTGGAAGCTGCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTTCTTGACACGACA 540
 QY 543 CCCATCAACAGTATATTTCTCCATGAAGACGGTACGGACTGGGCGTGGAGCATCTG 602
 DB 541 CCCATCAACAGTATATTTCTCCATGAAGACGGTACGGACTGGGCGTGGAGCATCTG 600
 QY 603 GTCGATTGGGTACACGAAATCGCGTGTAGCGGGCCCATTAAGTCTTGTTCTCGCG 662
 DB 601 GTCGATTGGGTACACGAAATCGCGTGTAGCGGGCCCATTAAGTCTTGTTCTCGCG 660
 QY 663 GCTCTGGCTCTGGCTGGCTGGCATATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
 DB 661 GCTCTGGCTCTGGCTGGCTGGCATATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 720
 QY 723 GAACGGGAAGCGCACTGAGTGGCATCTCGGTTTCAACAACCATGCAAAATGCTGAAT 782
 DB 721 GAACGGGAAGCGCACTGAGTGGCATCTCGGTTTCAACAACCATGCAAAATGCTGAAT 778
 QY 783 GAGG 786
 DB 779 GAGG 782

RESULT 5
 LOCUS B0751220
 DEFINITION B0751220 DSCT Colletotrichum trifolii cDNA clone pDST3-64, mRNA sequence.
 ACCESSION B0751220.1 GI:21906625
 VERSION B0751220
 KEYWORDS EST
 SOURCE Colletotrichum trifolii
 ORGANISM Colletotrichum trifolii
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes; Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae; mitosporic Phyllachoraceae; Colletotrichum.
 REFERENCE 1 (bases 1 to 806)
 AUTHORS Samac, D.A., Dickman, M., Town, C.D., Van Aken, S., Utterback, T., Cheung, F. and Fraser, C.M.
 TITLE ESTs from mycelia of Colletotrichum trifolii race 1
 JOURNAL Unpublished (2002)
 COMMENT Other ESTs: EST631782
 Contact: Deborah A. Samac
 Department of Plant Pathology
 University of Minnesota
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
 Tel: 612 625 1243
 Fax: 651 649 5058
 Email: debbys@pucini.crl.umn.edu
 TIGR sequence name: MTSAC64TV More information is available at:
 www.medicago.org
 Seq primer: (gtA AtA CgA Ctc Act AtA 99g C).
 Location/Qualifiers

FEATURES

source

1. .806
 /organism="Colletotrichum trifolii"
 /mol_type="mRNA"
 /strain="race 1"
 /db_xref="taxon:5466"
 /clone="pDST3-64"
 /tissue_type="mycelia"
 /dev_stage="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
 /lab_host="DH5alpha"
 /clone_lib="DST"
 /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2sp2 ; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gt11 from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid


```
/mol_type="genomic DNA"
/db_xref="taxon:186617"
/clone="SI051p36L"
/clone_lib="Scripps Pier (La Jolla, CA) uncultured virus
community"
/notes="Marine viruses were isolated from 200 liters of
surface seawater using a combination of differential
filtration and density-dependent gradient centrifugation.
Linker-amplified shotgun libraries were constructed by
randomly shearing the total marine viral community DNA,
end-repairing, ligating dsDNA linkers to the ends, and
amplifying the fragments using Vent DNA polymerase. The
resulting fragments were ligated into the pSMART vector
and electroporated into MC12 cells (Lucigen; Middleton,
WI)."

ORIGIN
Query Match 20.5%; Score 676; DB 8; Length 710;
Best Local Similarity 98.9%; Pred. No. 6.9e-189; Indels 1; Gaps 1;
Matches 690; Conservative 0; Mismatches 7; Indels 1; Gaps 1;
Qy 400 GTAAGCGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGTGATCAATTAATCAATC 459
Db 8 GTGCTGCAGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGTGATCAATTAATCAATC 67
Qy 460 CGCTGGATGACAGATGCCATTGCTGTGGAAGTGCCTGCACATTAATGTTCCGGCGTTAT 519
Db 68 CGCTGGATGACAGATGCCATTGCTGTGGAAGTGCCTGCACATTAATGTTCCGGCGTTAT 127
Qy 520 TTTCTTGATGTCCTGACAGACACCCATCAACAGTATTATTTTCCTCCATGAAGACGCTA 579
Db 128 TTTCTTGATGTCCTGACAGACACCCATCAACAGTATTATTTTCCTCCATGAAGACGCTA 187
Qy 580 CGCAGCTGGCGGTGAGCATCTGTGCGATTGGGTCAACAGCAATCGCGTGTGTAGCGG 639
Db 188 CGCAGCTGGCGGTGAGCATCTGTGCGATTGGGTCAACAGCAATCGCGTGTGTAGCGG 247
Qy 640 GCCATTAAAGTTCTGTCTCGGCGGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTC 699
Db 248 GCCATTAAAGTTCTGTCTCGGCGGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTC 307
Qy 700 GCAATCAATTCAGCCGATAGCGGAACGGGAAGCGACTGGAGTGCCATGTCGGGTTTC 759
Db 308 GCAATCAATTCAGCCGATAGCGGAACGGGAAGCGACTGGAGTGCCATGTCGGGTTTC 367
Qy 760 AACAAACCATGCAATGCTGAATGAGGGCATCGTTCCCACTGCCATGCTGGTTGCCAACG 819
Db 368 AACAAACCATGCAATGCTGAATGAGGGCATCGTTCCCACTGCCATGCTGGTTGCCAACG 427
Qy 820 ATCAGATGGCGTGGCGCAATGCGCGCATTAACGAGTCCGGGTGCGCGTGGTGGCGG 879
Db 428 ATCAGATGGCGTGGCGCAATGCGCGCATTAACGAGTCCGGGTGCGCGTGGTGGCGG 487
Qy 880 ATATCTCGGTAGTGGATACGACGATACCGAGACAGCTCATGTTATATCCCGCGTTAA 939
Db 488 ATATCTCGGTAGTGGATACGACGATACCGAGACAGCTCATGTTATATCCCGCGTTAA 547
Qy 940 CCACCATCAACAGAGATTTTCCTGCTGGGCAACACAGCGTGGACCGCTTGTGCAAC 999
Db 548 CCACCATCAACAGAGATTTTCCTGCTGGGCAACACAGCGTGGACCGCTTGTGCAAC 607
Qy 1000 TCTCTCAGGGCAGCGGTGAAGGCAATCAAGCTGTTGCGCGTCTCACTGGTGAAGAA 1059
Db 608 TCTCTCAGGGCAGCGGTGAAGGCAATCAAGCTGTTGCGCGTCTCACTGGTGAAGAA-AA 666
Qy 1060 AAACCAACCTGGCGCCCAATACGCAACCGCTCTCC 1097
Db 667 AAACCAACCTGGCGCCCAATACGCAACCGCTCTCC 704
```

RESULT 9
AL044483/c 660 bp mRNA linear EST 04-SEP-2003
LOCUS

```
DEFINITION DKFZp434i1102_s1 434 (synonym: htes3) Homo sapiens cDNA clone
AL044483 DKFZp434i1102 3', mRNA sequence.
VERSION AL044483.1 GI:5432701
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 660)
AUTHORS Ansorge,W., Benes,V., Krieger,S., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
TITLE EST (Ansorge, Benes, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 3' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No r1 sequence available.
This clone (DKFZp434i1102) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
LOCATION/Qualifiers
1..660
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp434i1102"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN
Query Match 19.6%; Score 647.4; DB 1; Length 660;
Best Local Similarity 99.7%; Pred. No. 2.2e-180;
Matches 659; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
Qy 85 AACCAAGTAACGTTATACGATCTCGCAGATGTCGGGTGCTCTTTATCAGACCGTTCC 144
Db 660 AACCAAGTAACGTTATACGATCTCGCAGATGTCGGGTGCTCTTTATCAGACCGTTCC 601
Qy 145 CGGTGTTGAACAGCCAGCCAGCCAGCTTTCTGCGAAACCGGGGAAAAGTGGNAGCGGCA 204
Db 600 CGGTGTTGAACAGCCAGCCAGCCAGCTTTCTGCGAAACCGGGGAAAAGTGGNAGCGGCA 541
Qy 205 TGGCGGAGCTGAATTACATTTCCCAACCGCTGGCAACAACCTGGCGGGCAACAGTGT 264
Db 540 TGGCGGAGCTGAATTACATTTCCCAACCGCTGGCAACAACCTGGCGGGCAACAGTGT 481
Qy 265 TGCTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGTCACGCGCGCTCGCAAAATTTGCGGG 324
Db 480 TGCTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGTCACGCGCGCTCGCAAAATTTGCGGG 421
Qy 325 CGATTAAATCTCGCGCCGATCAACTGGGTGTCAGGTTGGTGTGTCGATGTTAGAACGA 384
Db 420 CGATTAAATCTCGCGCCGATCAACTGGGTGTCAGGTTGGTGTGTCGATGTTAGAACGA 361
Qy 385 CGGGCTCGAAGCTGTAAAGCGGGGTGCACATCTTCTCGCGCAACCGCTCAGTGGGC 444
Db 360 CGGGCTCGAAGCTGTAAAGCGGGGTGCACATCTTCTCGCGCAACCGGTCACTGGGC 301
Qy 445 TGATCAATTAATCTCGCTGGATGACCAAGATGTCATTTGTTGGAAGCTGCTGCACCTA 504
Db 300 TGATCAATTAATCTCGCTGGATGACCAAGATGTCATTTGTTGGAAGCTGCTGCACCTA 241
Qy 505 ATGTTCCGGGTTATTTCTTGATGTCTCTGACCAACCCATCAACAGTATTATTTTCT 564
```



```

Db      240 ATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCAGACCCATCAACGATTAATTTTCT 181
QY      565 CCCATGAAGACGGTAGCGCACTGGCGTGGAGCATCTGTCGATTTGGGTACACGACAA 624
Db      180 CCCATGAAGACGGTAGCGCACTGGCGTGGAGCATCTGTCGATTTGGGTACACGACAA 121
QY      625 TCGCGTGTAGCGGCCCAATTAAGTTCTGTCTCGCGCGGTCTGCGTCTGGCTGGCTGGC 684
Db      120 TCGCGTGTAGCGGCCCAATTAAGTTCTGTCTCGCGCGGTCTGCGTCTGGCTGGCTGGC 62
QY      685 ATAAATATCTCACTCGCAATCAAAATTCACCGCATAGCGGAAACGGGAGCGCACTGGAGTG 744
Db      61 ATAAATATCTCACTCGCAATCAAAATTCACCGCATAGCGGAAACGGGAGCGCACTGGAGTG 744
QY      745 C 745
Db      1 C 1

RESULT 10
BJ063872/c
LOCUS
DEFINITION
Xenopus laevis (African clawed frog)
ACCESSION
BJ063872
VERSION
BJ063872.1
KEYWORDS
GI:17471074
SOURCE
Xenopus laevis
ORGANISM
Xenopus laevis (African clawed frog)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Xenopus.
REFERENCE
1 (bases 1 to 641)
Kitayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-i, T. and
Kohara, Y.
Expressed genes in X. laevis embryo
Unpublished (2001)
Contact: Tadao Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tehinigenes.nig.ac.jp
The information of this clone is available through the following
URL.
http://xenopus.nibb.ac.jp.
Location/Qualifiers
1. .641
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="XL07K02"
/tissue_type="whole embryo"
/dev stage="stage 25"
/clone_lib="NIBB Mochii normalized Xenopus tailbud
library"

FEATURES
source
Query Match 19.4%; Score 641; DB 4; Length 641;
Best Local Similarity 100.0%; Pred. NO. 1.7e-178;
Matches 641; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      30 CGCGGTATGCATGATACGCCCGGAGAGAGATCAATTCAGGTGGTGAATGTGAACCA 89
Db      641 CGCGGTATGCATGATACGCCCGGAGAGAGATCAATTCAGGTGGTGAATGTGAACCA 582
QY      90 GTAAACGTTATACGATGTGCGAGAGTAGTCCGGTGTCTCTTATCAGACCGTTTCCCGCGTG 149
Db      581 GTAAACGTTATACGATGTGCGAGAGTAGTCCGGTGTCTCTTATCAGACCGTTTCCCGCGTG 522
QY      150 GTGAACACGCCAGCCACGTTTCTCGGAAAAACCGGGAAAAAGTGAAGCGCGATGGCG 209

```

```

Db      521 GTGAACACGCCAGCCACGTTTCTCGGAAAAACCGGGAAAAAGTGAAGCGCGATGGCG 462
QY      210 GAGCTGAATTACATTCCCAACCGCGTGGCAACAACTGGCGGGCAACAGTCGTTGCTG 269
Db      461 GAGCTGAATTACATTCCCAACCGCGTGGCAACAACTGGCGGGCAACAGTCGTTGCTG 402
QY      270 ATTGGCGTTGCACTCCAGTCTGGCCCTGCAACGCGCGCTGCGAAATTTGTCGCGCGGAT 329
Db      401 ATTGGCGTTGCACTCCAGTCTGGCCCTGCAACGCGCGCTGCGAAATTTGTCGCGCGGAT 342
QY      330 AAATCTCGCGCGGATCAACTGGGTGCGAGCGTGGTGTGATGTTAGTAGAACGAGCGGC 389
Db      341 AAATCTCGCGCGGATCAACTGGGTGCGAGCGTGGTGTGATGTTAGTAGAACGAGCGGC 282
QY      390 GTCGAAGCCTGTAAAGCGGGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATC 449
Db      281 GTCGAAGCCTGTAAAGCGGGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATC 222
QY      450 ATTAACTATCGCTGGATGACCAAGATGCCATTTGCTGTGGAAGCTGCTGCACCTAATGTT 509
Db      221 ATTAACTATCGCTGGATGACCAAGATGCCATTTGCTGTGGAAGCTGCTGCACCTAATGTT 162
QY      510 CCGCGCTTATTTCTTGATGTCTCTGACCAACACCCATCAACAGTATTATTTTCTCCCAT 569
Db      161 CCGCGCTTATTTCTTGATGTCTCTGACCAACACCCATCAACAGTATTATTTTCTCCCAT 102
QY      570 GAAGACGTTACGGCACTGGCGGTGGAGCATCTCGTCGATTTGGTCCACGAGCAATTCGCG 629
Db      101 GAAGACGTTACGGCACTGGCGGTGGAGCATCTCGTCGATTTGGTCCACGAGCAATTCGCG 42
QY      630 CTGTTAGCGGGCCCATTAAGTTCTGTCGCGCGCTGCGGCGCTGCGG 670
Db      41 CTGTTAGCGGGCCCATTAAGTTCTGTCGCGCGCTGCTGCG 1

RESULT 11
CB863814/c
LOCUS
DEFINITION
HH07B20y HH Hordeum vulgare cDNA clone HH07B20 3-PRIME, mRNA
sequence.
ACCESSION
CB863814
VERSION
CB863814.1
KEYWORDS
GI:30058373
SOURCE
Hordeum vulgare
ORGANISM
Hordeum vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooidae; Triticeae; Hordeum.
1 (bases 1 to 624)
Varshney, R.K., Zhang, H., Burton, R., Stein, N., Langridge, P. and
Graner, A.
Barley ESTs from coleoptile tissue
Unpublished (2003)
Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 624 Std Error: 0.00
Plate: 7 row: B column: 20
Seq primer: SP6.
Location/Qualifiers
1. .624
/organism="Hordeum vulgare"
/mol_type="mRNA"
/db_xref="GABI:555471"
/db_xref="taxon:4513"
/clone="HH07B20"
/tissue_type="coleoptile"

```

ORIGIN	/dev_stage="coleoptile, 1 day old"	
	/lab_host="DH10B"	
TITLE	EST (Duesterhoeft, et al.)	
	Unpublished (1999)	
JOURNAL	CONTACT: MIPS	
	MIPS	
COMMENT	Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany	
	This is the 3' sequence of the clone insert	
FEATURES	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer	
	Research Center (DKFZ); sequenced by Qiagen within the cDNA	
source	sequencing consortium of the German Genome Project. r1 sequence	
	also available.	
1. .794	This clone is available at the RZPD in Berlin.	
	Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059	
/organism="Homo sapiens"	Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.	
	Location/Qualifiers	
/mol_type="mRNA"	1. .794	
	/db_xref="taxon:9606"	
/clone="DKFZp434B075"	/tissue_type="testis"	
	/dev_stage="adult"	
/lab_host="DH10B"	/lab_stage="adult"	
	/clone_lib="434 (synonym: htes3)"	
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"	/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"	
	ORIGIN	
Query Match	18.5%; Score 612; DB 1; Length 794;	
	Best Local Similarity 99.7%; Pred. No. 7.6e-170;	
Matches 617; Conservative 0; Mismatches 0; Indels 2; Gaps 2;	Matches 634; Conservative 0; Mismatches 0; Indels 2; Gaps 2;	
QY	569	TGAAGACGGTACGCGACTGGGGGTGGAGCATCTGTCGCATTTGGTTCACACGCAATCGC 628
Db	794	TGAAGACGGTACGCGACTGGGGGTGGAG-ATCTGTCGCAATTTGGTTCACACGCAATCGC 736
QY	629	GCTGTTAGCGGCGCCATTAAAGTTCTGCTCGGCGCGTCTGCGTCTGGCTGGCTGGCATAA 688
Db	735	GCTGTTAGCGGCGCCATTAAAGTTCTGCTCGGCGCGTCTG-GTCTGGCTGGCTGGCATAA 677
QY	699	ATATCTCACTCGCAATCAAAATTCAGCCGATAGCGAAACGGGAAGCGACTGAGATGCCAT 748
Db	676	ATATCTCACTCGCAATCAAAATTCAGCCGATAGCGAAACGGGAAGCGACTGAGATGCCAT 617
QY	749	GTCCGGTTTCAACAAACCATGCAATCTGAATCAGGGCATCGTTCCCATCTCCGATGCT 808
Db	616	GTCCGGTTTCAACAAACCATGCAATCTGAATCAGGGCATCGTTCCCATCTCCGATGCT 557
QY	809	GGTTCGCCAAGCATCAGATGGCGCTGGGCGCAATCGCGCATTTACCGAGTCCGGGCTGGC 868
Db	556	GGTTCGCCAAGCATCAGATGGCGCTGGGCGCAATCGCGCATTTACCGAGTCCGGGCTGGC 497
QY	869	CGTTGGTCCGATATCTCGGTAGTGGGATACGACATACCGAAGACAGCTCATGTTATAT 928
Db	496	CGTTGGTCCGATATCTCGGTAGTGGGATACGACATACCGAAGACAGCTCATGTTATAT 437
QY	929	CCGCGGTTTACCACCATCAAAACAGGATTTTCGCTCTGGGGGCAAAACGAGGTGACCG 988
Db	436	CCGCGGTTTACCACCATCAAAACAGGATTTTCGCTCTGGGGGCAAAACGAGGTGACCG 377
QY	989	CTTGTGCAACTCTCTCAGGCGCAGCGGTGAAGGCGCAATCAGCTGTTGCCCTCTCACT 1048
Db	376	CTTGTGCAACTCTCTCAGGCGCAGCGGTGAAGGCGCAATCAGCTGTTGCCCTCTCACT 317
QY	1049	GGTGAAGAAAGAAACCCACCTGGCGCCCAATAACGCAAAACCGCTCTCCCGCGCGTTGGC 1108
Db	316	GGTGAAGAAAGAAACCCACCTGGCGCCCAATAACGCAAAACCGCTCTCCCGCGCGTTGGC 257
QY	1109	CGATTCAATTAATGAGCTGGGACAGAGTTTCCCGACTGGAAAGCGGCGAGTGAAGCGCA 1168
Db	256	CGATTCAATTAATGAGCTGGGACAGAGTTTCCCGACTGGAAAGCGGCGAGTGAAGCGCA 197
QY	1169	ACGCAATTAATGAGTTAGCTCACTCACTATTAGGCAC 1204
Db	196	ACGCAATTAATGAGTTAGCTCACTCACTATTAGGCAC 161

ORIGIN	/dev_stage="coleoptile, 1 day old"	
	/lab_host="DH10B"	
TITLE	EST (Duesterhoeft, et al.)	
	Unpublished (1999)	
JOURNAL	CONTACT: MIPS	
	MIPS	
COMMENT	Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany	
	This is the 3' sequence of the clone insert	
FEATURES	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer	
	Research Center (DKFZ); sequenced by Qiagen within the cDNA	
source	sequencing consortium of the German Genome Project. r1 sequence	
	also available.	
1. .794	This clone is available at the RZPD in Berlin.	
	Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059	
/organism="Homo sapiens"	Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.	
	Location/Qualifiers	
/mol_type="mRNA"	1. .794	
	/db_xref="taxon:9606"	
/clone="DKFZp434B075"	/tissue_type="testis"	
	/dev_stage="adult"	
/lab_host="DH10B"	/lab_stage="adult"	
	/clone_lib="434 (synonym: htes3)"	
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"	/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"	
	ORIGIN	
Query Match	18.7%; Score 615.8; DB 6; Length 624;	
	Best Local Similarity 99.7%; Pred. No. 5.3e-171;	
Matches 617; Conservative 2; Mismatches 0; Indels 0; Gaps 0;	Matches 617; Conservative 2; Mismatches 0; Indels 0; Gaps 0;	
QY	3	GACACATCGATGTCGAAACCTTCGCGTATGGCATGTAGCGCCGGAAGAGT 62
Db	619	GACACATCGATGTCGAAACCTTCGCGTATGGCATGTAGCGCCGGAAGAGT 560
QY	63	CAATTACAGGTTGTAATGTAACACAGTAACGTTATACGATGTCGAGATGCGCGT 122
Db	559	CAATTACAGGTTGTAATGTAACACAGTAACGTTATACGATGTCGAGATGCGCGT 500
QY	123	GTCTCTTATCAGACCGTTTCCCGGTGTAACACAGGCGCAGCCAGCTTTCGCAAAACG 182
Db	499	GTCTCTTATCAGACCGTTTCCCGGTGTAACACAGGCGCAGCCAGCTTTCGCAAAACG 440
QY	183	CGGGAAGAGTCGAAGCGCGATGCGGAGCTGAATTACATTTCCCAACCGCGTGGCACA 242
Db	439	CGGGAAGAGTCGAAGCGCGATGCGGAGCTGAATTACATTTCCCAACCGCGTGGCACA 380
QY	243	CAACTGGCGGCAACACAGTCGTTGCTGATTGGCTTGGCCACCTCCAGTCTGCCCTGCAC 302
Db	379	CAACTGGCGGCAACACAGTCGTTGCTGATTGGCTTGGCCACCTCCAGTCTGCCCTGCAC 320
QY	303	GGCGCGTCGCAAAATGTTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCGTG 362
Db	319	GGCGCGTCGCAAAATGTTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCGTG 260
QY	363	GTGGTGTGATGTTAGAACGCGCGCTGGAAGCCTGTAAAGCGCGGTGCACAACTCTT 422
Db	259	GTGGTGTGATGTTAGAACGCGCGCTGGAAGCCTGTAAAGCGCGGTGCACAACTCTT 200
QY	423	CTCGCGCAACCGGTTCAGTGGGCTGATCAATTAATCTCGCTCGGTGACGAGATGCCATT 482
Db	199	CTCGCGCAACCGGTTCAGTGGGCTGATCAATTAATCTCGCTCGGTGACGAGATGCCATT 140
QY	483	GCTGTGGAAGCTGCGTGCATTAATGTTCCGGGTTATTTCTTGATGTTCTCTGACGAGCA 542
Db	139	GCTGTGGAAGCTGCGTGCATTAATGTTCCGGGTTATTTCTTGATGTTCTCTGACGAGCA 80
QY	543	CCCATCAACAGTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGGTGAGCATCTG 602
Db	79	CCCATCAACAGTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGGTGAGCATCTG 20
QY	603	GTCCGATTTGGTTCACCAGC 621
Db	19	GTCCGATTTGGTTCACCAGC 1

RESULT 12	AL045353	794 bp	mRNA	linear	EST 04-SEP-2003
LOCUS	DKFZp434B075.e1.434 (synonym: htes3)				Homo sapiens cDNA clone
DEFINITION	DKFZp434B075.3', mRNA sequence.				
ACCESSION	AL045353				
VERSION	AL045353.1	GI:5433506			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 794)				
AUTHORS	Duesterhoeft, A., Lauber, J., Mewes, H.W., Gassenhuber, J. and Wiemann, S.				

2414	AGACAAACCGCTGGGTGGGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATC	2475		
332	AGGCCAACCCGTTGGGTGGGTAGCGCTGAAGTCTTACGAGGAAGAGTTGGCGAAAGATC	273		
2474	CAGCTATTCCCGCCACCATGGAAACGCCAGAAAGGTGAATCATCGGAACATCCCGC	2533		
272	CAGCTATTCCCGCCACCATGGAAACGCCAGAAAGGTGAATCATCGGAACATCCCGC	213		
2534	AGATGTCGCTTTCTGTTATGCCGTGCTACTGCGGTGATCAACGCCCGACGCGTCTGTC	2593		
212	AGATGTCGCTTTCTGTTATGCCGTGCTACTGCGGTGATCAACGCCCGACGCGTCTGTC	153		
2594	AGACTGTGATGAAGCCCTGAAGAGCGGCAGACT	2628		
152	AGACTGTGATGAAGCCCTGAAGAGCGGCAGACT	118		
AY080094	608 bp	DNA	linear	GSS 06-NOV-2002
AY080094	Scripps Pier (La Jolla, CA)	uncultured virus community		
AY080094	uncultured marine virus genomic clone SIO5ip3d4L, genomic survey			
AY080094	sequence.			
AY080094	GI:24745276			
GSS.				
uncultured marine virus				
uncultured marine virus				
viruses; environmental samples.				
1 (bases 1 to 608)				
Breitbart,M., Salamon,P., Andresen,B., Mahaffy,J.M., Segall,A.M.,				
Mead,D., Azam,F. and Rohwer,F.				
Genomic analysis of uncultured marine viral communities				
Proc. Natl. Acad. Sci. U.S.A. 99 (22), 14250-14255 (2002)				
22294988				
12384570				
Contact: Rohwer F				
Biology Dept.				
San Diego State University				
5500 Campanile Dr, San Diego, CA 92102, USA				
Tel: 6195941336				
Fax: 619595676				
Email: forest@unstroke.sdsu.edu				
Class: Shotgun.				
Location/Qualifiers				
1. 608				
/organism="uncultured marine virus"				
/mol_type="genomic DNA"				
/db_xref="taxon:186617"				
/clone="SIO5ip3d4L"				
/clone_lib="Scripps Pier (La Jolla, CA) uncultured virus community"				
/notes="Marine viruses were isolated from 200 liters of surface seawater using a combination of differential filtration and density-dependent gradient centrifugation. Linker-amplified shotgun libraries were constructed by randomly shearing the total marine viral community DNA, end-repairing, ligating dsDNA linkers to the ends, and amplifying the fragments using Vent DNA polymerase. The resulting fragments were ligated into the pSMART vector and electroporated into MC12 cells (Lucigen; Middleton, WI)"				
Query Match	17.9%	Score	589.8;	DB 8; Length 608;
Best Local Similarity	99.7%	Pred. No.	2.8e-163;	
Matches	591;	Conservative	0;	Mismatches 2; Indels 0; Gaps 0;
287	CAGCTCGGCCCTGCACGCGCGTTCGCGCAATTCGCGCGGATTAATCTCGCGCGCATCA	346		
608	CAGCTCGGCCCTGCACGCGCGTTCGCGCAATTCGCGCGGATTAATCTCGCGCGCATCA	549		

QY 347 ACTGGTCCACCGTGGTGTGCTGATGGTAGAAGCGCGCTCGAAGCTGTAAAGC 406
Db ACTGGTCCACCGTGGTGTGCTGATGGTAGAAGCGCGCTCGAAGCTGTAAAGC 489
QY 407 GCGGTGCACAACTCTTCGCGCAACCGGTCACTGGGCTGATCAATACTATCCGCTGGA 466
Db GCGGTGCACAACTCTTCGCGCAACCGGTCACTGGGCTGATCAATACTATCCGCTGGA 429
QY 467 TGACCAAGATGCCATGTGCTGGAAAGCTGCCCTGCACTAATGTTCCAGCGTTATTTCTTGA 526
Db TGACCAAGATGCCATGTGCTGGAAAGCTGCCCTGCACTAATGTTCCAGCGTTATTTCTTGA 369
QY 527 TGTCTCTGACCAAGACACCCATCAACAGTATTTCTCCCATGAAGCGGTACCGGACT 586
Db TGTCTCTGACCAAGACACCCATCAACAGTATTTCTCCCATGAAGCGGTACCGGACT 309
QY 587 GGGCGTGGAGCATCTGGTTCGATTCGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAT 646
Db GGGCGTGGAGCATCTGGTTCGATTCGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAT 249
QY 647 AAGTTCTGTCTGCGCGCTGCTGCTGCTGGTGGCATAAATATCTCACTCGCAATCA 706
Db AAGTTCTGTCTGCGCGCTGCTGCTGCTGGTGGCATAAATATCTCACTCGCAATCA 189
QY 707 AATTGAGCGGATAGCGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTCAACAAC 766
Db AATTGAGCGGATAGCGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTCAACAAC 129
QY 767 CATGCAATGCTGAATGAGGCGATGTTCCCACTCGCATGCTGGTTGCCAAGATCAGAT 826
Db CATGCAATGCTGAATGAGGCGATGTTCCCACTCGCATGCTGGTTGCCAAGATCAGAT 69
QY 827 GCGCTGGCGCAATGCGCGCATTTACCGAGTCCGGCTGCGCGTGGTGGCG 879
Db GCGCTGGCGCAATGCGCGCATTTACCGAGTCCGGCTGCGCGTGGTGGCG 16

RESULT 15
AL037742/c
LOCUS DKFZp564C157 r1 564 (synonym: hfbr2) Homo sapiens cDNA clone
DEFINITION DKFZp564C157 5', mRNA sequence.
ACCESSION AL037742
VERSION 1 (bases 1 to 589)
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 589)
AUTHORS Bloecker,H., Boecker,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
EST (Bloecker, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No sl sequence available.
This clone (DKFZp564C157) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..589
1. .589
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp564C157"

FEATURES
source

/tissue_type="brain"
/dev_stage="fetal"
/lab_host="X1-2blue"
/clone_lib="564 (synonym: hfbr2)"
/note="Vector: pAMP1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 17.8%; Score 586.4; DB 1; Length 589;
Best Local Similarity 99.7%; Pred. No. 2.8e-162;
Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 517 TATTTCTTTGATGCTCTCGACAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACG 576
Db TATTTCTTTGATGCTCTCGACAGACACCCATCAACAGTATTATTTTCTCCCATGAAGACG 530
QY 577 GTACGCGACTGCGGCGTGAGCATCTGGTCGATTGGGTCAACAGCAAAATCGCGCTGTAG 636
Db GTACGCGACTGCGGCGTGAGCATCTGGTCGATTGGGTCAACAGCAAAATCGCGCTGTAG 470
QY 637 CGGGCCCAATTAAGTTCTGTCTCGGGCGCTCTGCGTCTGGCTGGCATAAATATCTCA 696
Db CGGGCCCAATTAAGTTCTGTCTCGGGCGCTCTGCGTCTGGCTGGCATAAATATCTCA 410
QY 697 CTGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCATGTCGGTT 756
Db CTGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGGCGACTGGAGTGCATGTCGGTT 350
QY 757 TTCAACAAACCATGCAAAATGCTGAATGAGGGCATCGTTCCCACTCGCATGCTGGTTGCCA 816
Db TTCAACAAACCATGCAAAATGCTGAATGAGGGCATCGTTCCCACTCGCATGCTGGTTGCCA 290
QY 817 ACATACAGATGCGCTGGGCGCAATGCGGCGCATTTACCGAGTCCGGGCTGCGGCTGGTG 876
Db ACATACAGATGCGCTGGGCGCAATGCGGCGCATTTACCGAGTCCGGGCTGCGGCTGGTG 230
QY 877 CGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCGT 936
Db CGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCGT 170
QY 937 TAACCAACCATCAAAACAGGATTTTCGCTGCTGGGCAAAACAGCGTGGACCGCTTGTGTC 996
Db TAACCAACCATCAAAACAGGATTTTCGCTGCTGGGCAAAACAGCGTGGACCGCTTGTGTC 110
QY 997 AACTCTCTCAGGGCCAGCGCGTGAAGGGCAATCAGCTGTTGCCCGCTCTCACCTGGTGA 1056
Db AACTCTCTCAGGGCCAGCGCGTGAAGGGCAATCAGCTGTTGCCCGCTCTCACCTGGTGA 50
QY 1057 GAAAAACCAACCTGGCGCCCAATACGCAACCGCCTCTCCCGCGCGTT 1105
Db GAAAAACCAACCTGGCGCCCAATACGCAACCGCCTCTCCCGCGCGTT 1

Search completed: August 3, 2005, 02:02:54
Job time : 6745.2 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 20:22:50 ; Search time 9276 Seconds
(without alignments)
17238.272 Million cell updates/sec

Title: US-09-765-555B-17
Perfect score: 3300
Sequence: 1 ccgacacatgaatggtgc.....acgaagttccggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sv.*
- 13: gb_un.*
- 14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	3300	100.0	3300	6	AX202429 Sequence
2	3232.8	98.0	3300	6	AX202430 Sequence
3	3215.2	97.4	3300	6	AX202426 Sequence
4	3215.2	97.4	3300	6	AX202427 Sequence
5	3207.2	97.2	3300	6	AX202428 Sequence
6	2705.8	82.0	7373	6	AX284144 Sequence
7	2705.8	82.0	8101	6	AX172306 Sequence
8	2701	81.8	6648	6	AX377531 Sequence
9	2700	81.8	7475	6	A91965 Sequence 1
10	2700	81.8	7475	6	AR031992 Sequence
11	2700	81.8	7475	6	AR207294 Sequence
12	2698	81.8	9191	6	AX377532 Sequence
13	2676.2	81.1	6724	12	AF097412 Expressio
14	2636.8	79.9	6806	6	AX378208 Sequence
15	2631	79.7	6706	12	AF031088 Shuttle v
16	2598.2	78.7	6748	12	AF031813 Expressio
17	1609.4	48.8	5558	6	CQ794769 Sequence
18	1474	44.7	4700	6	CQ846980 Sequence
19	1462.4	44.3	4921	6	AX768201 Sequence

20	1462.4	44.3	4935	6	AX768200	Sequence
21	1462.4	44.3	4945	6	AX768202	Sequence
22	1462.4	44.3	4951	6	AX768203	Sequence
c	23	1432.4	43.4	6755	12	AF140577 Integrati
	24	1417.6	43.0	1511	6	AX431576 Sequence
	25	1417.6	43.0	4297	12	AX219685 Expressio
	26	1417.6	43.0	4359	12	AX219686 Expressio
	27	1417.6	43.0	4419	12	AX219688 Expressio
	28	1417.6	43.0	4855	12	AX219689 Expressio
	29	1417.6	43.0	4968	12	AX219687 Expressio
	30	1417.6	43.0	6956	12	AX219682 Shuttle e
	31	1417.6	43.0	7018	12	AX219683 Shuttle e
	32	1417.6	43.0	7509	12	AX219684 Shuttle e
	33	1400.8	42.4	4566	12	AF177933 Cloning v
	34	1400.8	42.4	4635	12	AF177932 Cloning v
	35	1400.8	42.4	6399	12	AX302759 Expressio
	36	1400.8	42.4	6587	12	AX289542 Expressio
	37	1362	41.3	5201	6	AR493834 Sequence
	38	1362	41.3	5201	6	AR493835 Sequence
	39	1329.2	40.3	4806	12	AF050464 Expressio
c	40	1238.4	37.5	4803	12	AY243506 Cloning v
c	41	1238.4	37.5	5024	6	CQ846977 Sequence
	42	1237.2	37.5	1922	6	AX478528 Sequence
	43	1204	36.5	4506	12	AB116367 Expressio
	44	1204	36.5	4557	6	E13740 DNA sequenc
	45	1204	36.5	4557	6	AR373251 Sequence

ALIGNMENTS

RESULT 1
AX202429
LOCUS AX202429 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 17 from Patent WO0152620.
ACCESSION AX202429
VERSION AX202429.1 GI:15392177
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 17 26-JUL-2001;
The Scripps Research Institute (US) ; SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES
source location/Qualifiers
1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-m4 and zinc finger protein ZFPm4"

ORIGIN	Query Match	100.0%;	Score 3300;	DB 6;	Length 3300;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 3300;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CGACACCATCGAATGGTGCAGAACCTTTTCGCGGTATGGCATGTAGCGCCCGGAAGAGA	60		
Db	1	CGACACCATCGAATGGTGCAGAACCTTTTCGCGGTATGGCATGTAGCGCCCGGAAGAGA	60		
Qy	61	GTCAATTCAGGTGTGATGTGAACACGTAACCTTATACGATTCGCAGAGTATGCCG	120		
Db	61	GTCAATTCAGGTGTGATGTGAACACGTAACCTTATACGATTCGCAGAGTATGCCG	120		
Qy	121	GTGTCTCTTATCAGACCGTTCCTCCGCGGTGTGAACAGCCAGCCACGCTTTCTGCGAAAA	180		
Db	121	GTGTCTCTTATCAGACCGTTCCTCCGCGGTGTGAACAGCCAGCCACGCTTTCTGCGAAAA	180		
Qy	181	CGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC	240		

181	Db	CGCGGGAAAAAGTGGAAAGCGCGGATGGCGGAGCTGAAATTACATTTCCCAACCGCGTGGCAC	240
241	Qy	AACAACATGGCGGGGCAACAGTCTGTTTCTGATTTGGCGTTTGCCACCTCCAGTCTGGCCCTGCG	300
241	Db	AACAACATGGCGGGGCAACAGTCTGTTTCTGATTTGGCGTTTGCCACCTCCAGTCTGGCCCTGCG	300
301	Qy	ACGCGCGGTGCGCAAAATTTGTTCGCGCGGATTAATCTCGCGCGGATCAACTGGTGGTCCAGCG	360
301	Db	ACGCGCGGTGCGCAAAATTTGTTCGCGCGGATTAATCTCGCGCGGATCAACTGGTGGTCCAGCG	360
361	Qy	TGTTGGTGTGCATGGTAGAACGAAGCGCGCTGGAAGCCTGTAAGCGCGCGGTGCACAAATC	420
361	Db	TGTTGGTGTGCATGGTAGAACGAAGCGCGCTGGAAGCCTGTAAGCGCGCGGTGCACAAATC	420
421	Qy	TTCTCGGCGAACGGGTCAGTGGGCTGATCATTAATCTATCGCTCGATGACACGAGTACCCA	480
421	Db	TTCTCGGCGAACGGGTCAGTGGGCTGATCATTAATCTATCGCTCGATGACACGAGTACCCA	480
481	Qy	TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTTGATGTCTCTGACACAGA	540
481	Db	TTGCTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTTATTTCTTTGATGTCTCTGACACAGA	540
541	Qy	CACCCATCAACAGTATTATTTTCTCCCATGGAAGAGGTACCGGATCTGGGCGTGGAGCATC	600
541	Db	CACCCATCAACAGTATTATTTTCTCCCATGGAAGAGGTACCGGATCTGGGCGTGGAGCATC	600
601	Qy	TGGTTCGCAATGGGTGACACAGCAAAATCGCGCTGTAGCGGGGCCAATTAAGTTCTCTCTCGG	660
601	Db	TGGTTCGCAATGGGTGACACAGCAAAATCGCGCTGTAGCGGGGCCAATTAAGTTCTCTCTCGG	660
661	Qy	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
661	Db	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
721	Qy	CGGAAACGGGAAAGCGCATGGAAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA	780
721	Db	CGGAAACGGGAAAGCGCATGGAAGTGCATGTCCGGTTTTCAACAAACCATGCAAAATGCTGA	780
781	Qy	ATGAGGGCATGTTTCCCACTCGGATGCTGTTGTCACAGATCAGATGGCGCTGGCGGCAAC	840
781	Db	ATGAGGGCATGTTTCCCACTCGGATGCTGTTGTCACAGATCAGATGGCGCTGGCGGCAAC	840
841	Qy	TGCGCGCATTACCGAGTCCGGGCTGCGCGTGTGTTGGGATATCTCGTGTAGTGGGATACG	900
841	Db	TGCGCGCATTACCGAGTCCGGGCTGCGCGTGTGTTGGGATATCTCGTGTAGTGGGATACG	900
901	Qy	ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTTAAACCAACATCAAAACAGGATTTTC	960
901	Db	ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTTAAACCAACATCAAAACAGGATTTTC	960
961	Qy	GCCTGTGGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
961	Db	GCCTGTGGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
1021	Qy	AGGGCAATCAGCTGTGTCGCCCTCTCACTGTGTGAAAGAAAAACCACTCTGGCGGCCAATA	1080
1021	Db	AGGGCAATCAGCTGTGTCGCCCTCTCACTGTGTGAAAGAAAAACCACTCTGGCGGCCAATA	1080
1081	Qy	CGCAAAACCGCTCTCTCCCGCGGTTGCGCGATTCAATTAATGCACTGGCAGCAGAGTTT	1140
1081	Db	CGCAAAACCGCTCTCTCCCGCGGTTGCGCGATTCAATTAATGCACTGGCAGCAGAGTTT	1140
1141	Qy	CCGACTGGAAAGCGGGCAGTGAAGGCAACCGAATTAATGTGAGTTAGCTCACTCATTTAG	1200
1141	Db	CCGACTGGAAAGCGGGCAGTGAAGGCAACCGAATTAATGTGAGTTAGCTCACTCATTTAG	1200
1201	Qy	GCAAAATTCATGTTTGACAGCTTATCATCGACTGACGCGTGACCAAAATGCTTCTGGCG	1260
1201	Db	GCAAAATTCATGTTTGACAGCTTATCATCGACTGACGCGTGACCAAAATGCTTCTGGCG	1260
1261	Qy	TCAGCGAGCCATCGGAAGCTGTGTTGCTGTGCAAGTGTGTAATCACTGCATAATTCG	1320

QY 841 TCGCGCCATTACCGAGTCCGGCTCGCGTGTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TCGCGCCATTACCGAGTCCGGCTCGCGTGTGGTGGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTATATATCCCGCGGTTAAACCAATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTATATATCCCGCGGTTAAACCAATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACAGCTGGACCGCTTCTGCTGCAAAAGAAAACCAACCTGCGCGCCCAATA 1020
Db 961 GCCTGCTGGGCAACACAGCTGGACCGCTTCTGCTGCAAAAGAAAACCAACCTGCGCGCCCAATA 1020
QY 1021 AGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAACCAACCTGCGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAACCAACCTGCGCGCCCAATA 1080
QY 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGCTGGCAGCAGGTTT 1140
Db 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCGAGCTGGCAGCAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACGGTGCAACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACGGTGCAACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTCGAGGTCGTAATCACTGCAATTTG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTCGAGGTCGTAATCACTGCAATTTG 1320
QY 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATTAATGTTTTTGGCGGACATCAACAGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATTAATGTTTTTGGCGGACATCAACAGGTT 1380
QY 1381 CTGGCAATATTCTGAATGAGCTGTGAACAATTAATCATCGGTCGGTATTAATGTGGA 1440
Db 1381 CTGGCAATATTCTGAATGAGCTGTGAACAATTAATCATCGGTCGGTATTAATGTGGA 1440
QY 1441 ATTGTAGCGGATAACAATTTACACAGGAACAGCGAGTCCGTTTAGTGTGTTTTCAGA 1500
Db 1441 ATTGTAGCGGATAACAATTTACACAGGAACAGCGAGTCCGTTTAGTGTGTTTTCAGA 1500
QY 1501 GCATTCACCAACAGGACCATAGATTATGAAACTGAAGAGGTAACTGGTAACTCTGG 1560
Db 1501 GCATTCACCAACAGGACCATAGATTATGAAACTGAAGAGGTAACTGGTAACTCTGG 1560
QY 1561 ATTAACGGCGATAAGGCTATAACGGTCTCGCTGAAGTGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAGGCTATAACGGTCTCGCTGAAGTGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCGGAAGAGAAATTCGAGAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCGGAAGAGAAATTCGAGAGGTT 1680
QY 1681 GCGGCAATCGCGATGCGCTGACATTATCTTGGGCAACAGCAGGCTTGGTGGCTAC 1740
Db 1681 GCGGCAATCGCGATGCGCTGACATTATCTTGGGCAACAGCAGGCTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCGTGTGGCTGAATCAACCGGACAAAGGTTCCAGGACAGCTGTAT 1800
Db 1741 GCTCAATCTGGCGTGTGGCTGAATCAACCGGACAAAGGTTCCAGGACAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGTACGTTTCAACGGCAGCTGATTGCTTACCGGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGTACGTTTCAACGGCAGCTGATTGCTTACCGGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGTCGCGAACCGCGCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGTCGCGAACCGCGCAAAACCTGGGAA 1920

QY 1921 GAGATCCGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTTCACCTGGCGCTGATTCCTGCTGACGGGGGTATGCGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGGCGCTGATTCCTGCTGACGGGGGTATGCGGTTCAAG 2040
QY 2041 TATGAAAAACCGCAAGTAGACGATTAAGAGAGCTGGCGCTGGATAAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAAACCGCAAGTAGACGATTAAGAGAGCTGGCGCTGGATTAAGAGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAAGAAACACATGATGACAGACCGGATAC 2160
Db 2101 GGTCTGACCTTCTGCTGTTGACCTGATTAAGAAACACATGATGACAGACCGGATAC 2160
QY 2161 TCCATCGAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGGCGCGCTGG 2220
Db 2161 TCCATCGAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGGCGCGCTGG 2220
QY 2221 GCATGCTCAACATCGACACAGCAAAAGTGAATTAATGCTGTAAACGGTACTGCGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACAGCAAAAGTGAATTAATGCTGTAAACGGTACTGCGACCTTC 2280
QY 2281 AAGGCTCAACATCCAAACCGTTTCTGTTGGCTGCTGAGCGCAGGTATTAACGGCGCCAGT 2340
Db 2281 AAGGCTCAACATCCAAACCGTTTCTGTTGGCTGCTGAGCGCAGGTATTAACGGCGCCAGT 2340
QY 2341 CCAGAACAAAGAGCTGGCAAAAGATTCCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
Db 2341 CCAGAACAAAGAGCTGGCAAAAGATTCCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTAGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTAGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCAGTATTGCGGCAACCATCGAAAAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTGCGGCAACCATCGAAAAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCAGAACATCCCGCAGATGCTCGTTTCTGTTGATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCAGAACATCCCGCAGATGCTCGTTTCTGTTGATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGCTGCTCAGACTGTGCGATGAAGCCCTGAAAGACGCGAGACTTAATTCGAGCTG 2640
Db 2581 GCCAGCGCTGCTCAGACTGTGCGATGAAGCCCTGAAAGACGCGAGACTTAATTCGAGCTG 2640
QY 2641 AACAAACAAACAAATAACAAATAACAAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACAAACAAACAAATAACAAATAACAAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
QY 2701 GGATCTCTTCTCTGTCGCCAGGCGGCTTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTGTCGCCAGGCGGCTTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTGGTAAGTCTTTCAGCCAGAGCAGCTCCCTGCTGCGCCACAGCGTATCCAC 2820
Db 2761 CCGGAATGTGGTAAGTCTTTCAGCCAGAGCAGCTCCCTGCTGCGCCACAGCGTATCCAC 2820
QY 2821 ACGGGTGAAGAACCGGTATTAATGCGGAGGAGCCATCAAAATGTCAGATGAGCAGCAGC 2880
Db 2821 ACGGGTGAAGAACCGGTATTAATGCGGAGGAGCCATCAAAATGTCAGATGAGCAGCAGC 2880
QY 2881 CTGGTCCGCCATCAACGCACTCATCTGCGGAGAGCCATCAAAATGTCAGATGAGCAGCAGC 2940
Db 2881 CTGGTCCGCCATCAACGCACTCATCTGCGGAGAGCCATCAAAATGTCAGATGAGCAGCAGC 2940
QY 2941 AAGTCTTTTCAAGTGTGCTGATCTTTCGAGGACCAACCGTACTCACACCGGGGAGAG 3000
Db 2941 AAGTCTTTTCAAGTGTGCTGATCTTTCGAGGACCAACCGTACTCACACCGGGGAGAG 3000
QY 3001 CCTATGCTTGTCCGGAATGTGGTAAGTCTTCTCTCAGAGCTCTCACCTGGTGGCCAC 3060

|||||
3001 CCCTATGCTTGCCTGGAATAGTGGTAAGTCTTTCAGCACCGAGTGGTCTCTTGGTTAGACAC 3060
QY
3061 CAGCGTACCCACACACGGGTGAAAAACCGTATATAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
Db
3061 CAGCGTACCCACACACGGGTGAAAAACCGTATATAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
QY
3121 CGCAGCGATACTCGTGGGCCATCAACGCACTATCTACTGGCGAGAACCATACAAATGT 3180
Db
3121 CAGCGCGCCACCTGGAAAGCCCATCAACGCACTATCTACTGGCGAGAACCATACAAATGT 3180
QY
3181 CCAGAAATGGCAAGTCTTCTCAACTTCAGGCCATTTGGTCCGTCAACACGTACTCAC 3240
Db
3181 CCAGAAATGGCAAGTCTTCTCAACTTCAGGCCAATTTGGTCCGTCAACACGTACTCAC 3240
QY
3241 ACCGGTAAAAAATAGTGGCCAGCGCGGCCAGTACCCGTACGAGTTCGGGACTACGCT 3300
Db
3241 ACCGGTAAAAAATAGTGGCCAGCGCGGCCAGTACCCGTACGAGTTCGGGACTACGCT 3300

RESULT 3
AX202426
LOCUS AX202426 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 14 from Patent WO0152620.
ACCESSION AX202426
VERSION AX202426.1 GI:15392173
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 14 26-JUL-2001;
The Scripps Research Institute (US) ; SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)

FEATURES
source Location/Qualifiers
1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-m1 and zinc finger protein
zFPml"

ORIGIN
Query Match 97.4%; Score 3215.2; DB 6; Length 3300;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 3247; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGAGAGTATGCCG 120
Db 61 GTCAATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGAGAGTATGCCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCTCCGCGTGTGAACACAGCCAGCCAGCTTTCTCGGAAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCTCCGCGTGTGAACACAGCCAGCCAGCTTTCTCGGAAAA 180
QY 181 CGCGGAAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAAAAGTGTGCTGATTTGGGCTGGCCAGCTTCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGCAAAAGTGTGCTGATTTGGGCTGGCCAGCTTCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTTCGCAAAATTTGTCGGGGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGTTCGCAAAATTTGTCGGGGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 360

QY 361 TGGTGGTGTGATCGTAGAACGAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCACAATC 420
Db 361 TGGTGGTGTGATCGTAGAACGAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCACAATC 420
QY 421 TTCTCGCGCAACAGCGTCACTAGTGGGCTGATCACTAACTATCCGCTGGATGACACAGGATGCCA 480
Db 421 TTCTCGCGCAACAGCGTCACTAGTGGGCTGATCACTAACTATCCGCTGGATGACACAGGATGCCA 480
QY 481 TTGCTGTGGAAAGTGCCTGCACTAATGTTCCGGGGTATTTCTTGATGTCCTGACCAAGA 540
Db 481 TTGCTGTGGAAAGTGCCTGCACTAATGTTCCGGGGTATTTCTTGATGTCCTGACCAAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACCGACTGGCGGTGGAGCATC 600
QY 601 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTGTAGCGGGGCCATTAAAGTGTCTGCTCGG 660
Db 601 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTGTAGCGGGGCCATTAAAGTGTCTGCTCGG 660
QY 661 CGCGTCTCGCTCGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTCGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCGACTGGAGTGGCATGTCCGGTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGCGACTGGAGTGGCATGTCCGGTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTCGCATGTGTGTGCCAAGCATCAGATGCGCTGGCGGCA 840
Db 781 ATGAGGGCATGTTCCCACTCGCATGTGTGTGCCAAGCATCAGATGCGCTGGCGGCA 840
QY 841 TCGCGCCCATTTACCGAGTCCGGGTGCGCGTGTGTGCGGATATCTCGGTAGTGGGATACG 900
Db 841 TCGCGCCCATTTACCGAGTCCGGGTGCGCGTGTGTGCGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAAACCAACCATCAACAGSATTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAAACCAACCATCAACAGSATTTC 960
QY 961 GCCTGCTGGGGCAACACAGCGTGGACCGCTTGTCTCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGGCAACACAGCGTGGACCGCTTGTCTCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTGCCCGCTCTCACTGGTGAAGAAAAAACCCCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGCCCGCTCTCACTGGTGAAGAAAAAACCCCTGGCGCCCAATA 1080
QY 1081 GCAGAACCGCTCTCCCGCGGTGGCGGATTCATTAATGCGAGCTGGCAGCAGAGTTT 1140
Db 1081 GCAGAACCGCTCTCCCGCGGTGGCGGATTCATTAATGCGAGCTGGCAGCAGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTTAGTCTACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTTAGTCTACTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCGGTGCAACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCGGTGCAACCAATGCTTCTGGCG 1260
QY 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTGAATTCATCTGATTAATTCG 1320
Db 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTGAATTCATCTGATTAATTCG 1320
QY 1321 TGTGCTCAAGCGCACTCCCGTTCTGGTAAATGTTTTTTCGCCCGACATCAACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTTCTGGTAAATGTTTTTTCGCCCGACATCAACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTGCACAATTAATCATCGGCTCGTATATGTGTGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTGCACAATTAATCATCGGCTCGTATATGTGTGA 1440
QY 1441 ATTGTGAGCGGATAACAATTTTTCACAGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500

```
Db 1441 ATTGTAGCGGATAACAATTTTACACAGGAACGCCAGTCGGTTAGGTGTTTTCAGCA 1500
Qy 1501 GCACCTTCCACCAAGGACCATAGATTATGAAACTGAAGAGTAAACTGTAATCTGG 1560
Db 1501 GCACCTTCCACCAAGGACCATAGATTATGAAACTGAAGAGTAAACTGTAATCTGG 1560
Qy 1561 ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCGGAAGAGAAATCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCGGAAGAGAAATCCACAGGTT 1680
Qy 1681 GCGGCAATCGGCGATGGCCCTGACATTATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAATCGGCGATGGCCCTGACATTATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCTGTTGGCTGAATTAACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAATTAACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGGCGTACGTTTACACCGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGGCGTACGTTTACACCGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
Qy 1861 GAAAGCTTATCGCTGATTATAACAAGATCTGCTGCCGAAACCCCGCAAAACCTGGGAA 1920
Db 1861 GAAAGCTTATCGCTGATTATAACAAGATCTGCTGCCGAAACCCCGCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAGCGAAAGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAGCGAAAGTAAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAACGGCAAGTACGACATTAAGACCTGGCGCTGGATAACGCTGGCGCAAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACCTGGCGCTGGATAACGCTGGCGCAAAAGCG 2100
Qy 2101 GGTCTGACCTTCCTGGTTGACCTGATTAATAAACAACAACATGAATGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCCTGGTTGACCTGATTAATAAACAACAACATGAATGAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCCAAATCGACACAGCAAAAGTGAATTAATGTTGTAACGGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACAGCAAAAGTGAATTAATGTTGTAACGGTACTGCCGACCTTC 2280
Qy 2281 AAGGGTCAACCATCAACCGTTCGTTGGCGTCTGAGCGCAGGTATTAACCGCCAGT 2340
Db 2281 AAGGGTCAACCATCAACCGTTCGTTGGCGTCTGAGCGCAGGTATTAACCGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTCTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTCTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAACGGTTTAATAAGCAAAACCGCTGGTGCCTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAACGGTTTAATAAGCAAAACCGCTGGTGCCTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGCGAAAGATCCACGTTATTCGCGCACCATGGAACCGCCCAAGAGTGAATCATG 2520
Db 2461 TTGCGAAAGATCCACGTTATTCGCGCACCATGGAACCGCCCAAGAGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGGCTTTCTGGTATGCCGTGGCTACTCGCGTGATCAACGCC 2580
```

```
Db 2521 CCGAACATCCCGCAGATGTCGGCTTTCTGGTATGTCGCTGCGTACTCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGATGATGATGAAGCCCTGAAAGACCGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGATGATGATGAAGCCCTGAAAGACCGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACAAACAACAATAAATAACAACAACCTCGGGATCGAGGAAGATTTTCAAGATTC 2700
Db 2641 AACAAACAACAATAAATAACAACAACCTCGGGATCGAGGAAGATTTTCAAGATTC 2700
Qy 2701 GGATCTCTTCTCTCTGTCGCGCCAGCGCGCTCTCGAGCCCGGGAGAGACCTATGCTTGT 2760
Db 2701 GGATCTCTTCTCTCTGTCGCGCCAGCGCGCTCTCGAGCCCGGGAGAGACCTATGCTTGT 2760
Qy 2761 CCGGAATGTTGTAAGTCTTCTCAGCAGAGCTCTCTCTCAGAGCTCTCACCTGGTGGCCAC 2820
Db 2761 CCGGAATGTTGTAAGTCTTCTCAGCAGAGCTCTCTCTCAGAGCTCTCACCTGGTGGCCAC 2820
Qy 2821 ACCGGTGAATAAACCCTATAAATGCCAGAGTGGGCAAAATCTTTTAGCCAGAGCAGCAGC 2880
Db 2821 ACCGGTGAATAAACCCTATAAATGCCAGAGTGGGCAAAATCTTTTAGCCAGTCCAGCAAC 2880
Qy 2881 CTGTCGCGCATCAACCGACTCATCTACTGGCGAGAAGCCATACAAATGTCAGAAATGTGGC 2940
Db 2881 CTGTCGCGCATCAACCGACTCATCTACTGGCGAGAAGCCATACAAATGTCAGAAATGTGGC 2940
Qy 2941 AAGTCTTTTCAAGTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 3000
Db 2941 AAGTCTTTTCAAGTCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 3000
Qy 3001 CCTATCTTCTGTCGGAATGTTGTAAGTCTTCTCTCAGAGCTCTCACCTGGTGGCGCAC 3060
Db 3001 CCTATCTTCTGTCGGAATGTTGTAAGTCTTCTCAGCGCAGCGATTAACCTGGTGGCGCAC 3060
Qy 3061 CAGGTACCCACACCGGGTGAATAAACCCTATAAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
Db 3061 CAGGTACCCACACCGGGTGAATAAACCCTATAAATGCCAGAGTGGCGCAAAATCTTTTAGC 3120
Qy 3121 CGCAGCGATTAACCTGGTGGCCCATCAACGACTCATCTGCGCAGAGCGCATTAACCTGTC 3180
Db 3121 CAGCGCGCCACTGGCCAGCCATCAACGACTCATCTGCGCAGAGCGCATTAACCTGTC 3180
Qy 3181 CCAGATGTCGCAAGTCTTTTCTCAACTTCAGGCGATTTGTCGCTCACCAACCGTACTCAC 3240
Db 3181 CCAGATGTCGCAAGTCTTTTCTCTCTGTCGCAAACTCTGTCGCGCAGCAAACTACTCAC 3240
Qy 3241 ACCGTAATAAATAACTAGTGGCCAGCGCGCGCAGTACCGTACGAGTTCGCGACTAGCT 3300
Db 3241 ACCGTAATAAATAACTAGTGGCCAGCGCGCGCAGTACCGTACGAGTTCGCGACTAGCT 3300
```

RESULT 4

AX202427 LOCUS AX202427 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 15 from Patent WO0152620.
ACCESSION AX202427
VERSION AX202427.1 GI:15392175
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Barbas, C.F., Stege, J.T., Guan, X. and Dalmia, B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 15 26-JUL-2001;
The Scripps Research Institute (US); SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES
Location/Qualifiers
1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-m2 and zinc finger protein"

ORIGIN		2ZPm2"	
Query Match		97.4%; Score 3215.2; DB 6; Length 3300;	
Best Local Similarity		98.4%; Pred. No. 0;	
Matches 3247; Conservative		0; Mismatches 53; Indels 0; Gaps 0;	
QY	1	CCGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA	60
DB	1	CCGACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA	60
QY	61	GTCAATTACAGGTGGTGAATGTGAACACAGTAACAGTTATACGATGTCGCAGAGTATGCG	120
DB	61	GTCAATTACAGGTGGTGAATGTGAACACAGTAACAGTTATACGATGTCGCAGAGTATGCG	120
QY	121	GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCACGTTTCTCGGAAAA	180
DB	121	GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCACGTTTCTCGGAAAA	180
QY	181	CGCGGAAAAAGTGAAGCGGCGATGGCGAGCTGAAATTAATCCCAACCGCGTGGCAC	240
DB	181	CGCGGAAAAAGTGAAGCGGCGATGGCGAGCTGAAATTAATCCCAACCGCGTGGCAC	240
QY	241	AACAACTGCGCGCAACAGTCGTTGCTGATTTGGCGTTTGCACCTCCAGTCTGGCCCTGC	300
DB	241	AACAACTGCGCGCAACAGTCGTTGCTGATTTGGCGTTTGCACCTCCAGTCTGGCCCTGC	300
QY	301	AGCGCCGTTCGCAAAATTTGTTCGCGCGATTAATAATCTCGCCCGATCAACTGGGTGCCAGC	360
DB	301	AGCGCCGTTCGCAAAATTTGTTCGCGCGATTAATAATCTCGCCCGATCAACTGGGTGCCAGC	360
QY	361	TGGTGGTTCGATGTAGAACGAAAGCGGCTGAAAGCCCTGTAAGCGGCGTGCACAAATC	420
DB	361	TGGTGGTTCGATGTAGAACGAAAGCGGCTGAAAGCCCTGTAAGCGGCGTGCACAAATC	420
QY	421	TTCTCGCGCAACGCTGAGTGGGCTGATCATTAATCTATCCGCTGATGACACAGGATGCCA	480
DB	421	TTCTCGCGCAACGCTGAGTGGGCTGATCATTAATCTATCCGCTGATGACACAGGATGCCA	480
QY	481	TTGCTGTGGAAGCTGCCTGCACTAATGTTTCGGCGTTATTTCTTGATGTTCTCTGACCAGA	540
DB	481	TTGCTGTGGAAGCTGCCTGCACTAATGTTTCGGCGTTATTTCTTGATGTTCTCTGACCAGA	540
QY	541	CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600
DB	541	CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600
QY	601	TGGTGCATTTGGTCCACGAAATCGCGCTGTAGCGGCCCATTTAAGTTCTGTCTCGG	660
DB	601	TGGTGCATTTGGTCCACGAAATCGCGCTGTAGCGGCCCATTTAAGTTCTGTCTCGG	660
QY	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAG	720
DB	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAATTCAGCCGATAG	720
QY	721	CGGAAACGGGAAGCGACTGGAAGTGCATGTCGCGTTTCAACAAACCATGCAAAATGCTGA	780
DB	721	CGGAAACGGGAAGCGACTGGAAGTGCATGTCGCGTTTCAACAAACCATGCAAAATGCTGA	780
QY	781	ATGAGGGCATCGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGGCGCTGGGGGCA	840
DB	781	ATGAGGGCATCGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGGCGCTGGGGGCA	840
QY	841	TGCGGCGCAATTACAGATCGGGCTGCGGTTGGTGGCATATCTCGGTAGTGGGATAG	900
DB	841	TGCGGCGCAATTACAGATCGGGCTGCGGTTGGTGGCATATCTCGGTAGTGGGATAG	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCACTCAACAGGATTTTC	960
DB	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCACTCAACAGGATTTTC	960
QY	961	GCCTGCTGGGGCAACACGCGTGGACCGCTTCTGCAACTCTCAGGCGCAGGCGGTGA	1020
DB	961	GCCTGCTGGGGCAACACGCGTGGACCGCTTCTGCAACTCTCAGGCGCAGGCGGTGA	1020

DB	961	GCCTGCTGGGGCAACACGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCAGGCGGTGA	1020
QY	1021	AGGCAATCAGCTGTGTCCCTCTCTCACTGGTGAAAGAAAAACCACTCTGGCGCCCAATA	1080
DB	1021	AGGCAATCAGCTGTGTCCCTCTCTCACTGGTGAAAGAAAAACCACTCTGGCGCCCAATA	1080
QY	1081	CGAAGACCGCTCTCCCGCGGTTGGCCGATTCATTAATGCAGCTGGCAGCAGAGTTT	1140
DB	1081	CGAAGACCGCTCTCCCGCGGTTGGCCGATTCATTAATGCAGCTGGCAGCAGAGTTT	1140
QY	1141	CCGACTCGAAGAGGGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200
DB	1141	CCGACTCGAAGAGGGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG	1200
QY	1201	GCACAAATTCATGTTTGACAGCTTATCATCGACTGCGCGGTGCAACCAATGCTTCTGGCG	1260
DB	1201	GCACAAATTCATGTTTGACAGCTTATCATCGACTGCGCGGTGCAACCAATGCTTCTGGCG	1260
QY	1261	TCAGCAGCCATCGGAGCTGTGGTATGGCTGTGCAGGTGTAATCACTGCATTAATTCG	1320
DB	1261	TCAGCAGCCATCGGAGCTGTGGTATGGCTGTGCAGGTGTAATCACTGCATTAATTCG	1320
QY	1321	TGTGCTCAAGCGCACCTCCCGTCTTGATTAATGTTTTTGGCGCGACATCATACCGTT	1380
DB	1321	TGTGCTCAAGCGCACCTCCCGTCTTGATTAATGTTTTTGGCGCGACATCATACCGTT	1380
QY	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGGA	1440
DB	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGGA	1440
QY	1441	ATTGTGAGCGGATTAACAAATTTTACACAGGAAACGCGAGTCCGTTTAGGTGTTTTCAGA	1500
DB	1441	ATTGTGAGCGGATTAACAAATTTTACACAGGAAACGCGAGTCCGTTTAGGTGTTTTCAGA	1500
QY	1501	GCACCTTCAACCAAGGACCATAGATTATGAAACTGAAGAGGTAAACTGGTAAATCTGG	1560
DB	1501	GCACCTTCAACCAAGGACCATAGATTATGAAACTGAAGAGGTAAACTGGTAAATCTGG	1560
QY	1561	ATTAAACGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT	1620
DB	1561	ATTAAACGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAAGTACACCGTTGAGCATCCGATTAACCTGGAGAGAGAAATCCCAAGGTT	1680
DB	1621	ACCGGAATTAAGTACACCGTTGAGCATCCGATTAACCTGGAGAGAGAAATCCCAAGGTT	1680
QY	1681	CGCGCAACTGCGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC	1740
DB	1681	CGCGCAACTGCGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCTGTGGCTGAAATCACCCGGAACAAAGCGTTCCAGGACAAGCTGTAT	1800
DB	1741	GCTCAATCTGGCTGTGGCTGAAATCACCCGGAACAAAGCGTTCCAGGACAAGCTGTAT	1800
QY	1801	CGGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTCCTTACCGATTCGTT	1860
DB	1801	CGGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTCCTTACCGATTCGTT	1860
QY	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCCGCAAAACCTGGGAA	1920
DB	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCCGCAAAACCTGGGAA	1920
QY	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGGTGAAGAGCGCGCTGATGTTCAAC	1980
DB	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGGTGAAGAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATGTTGCTGACCGGGGTTATCGGTTCAAG	2040
DB	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATGTTGCTGACCGGGGTTATCGGTTCAAG	2040
QY	2041	TATGAAAACGCGCAAGTACGACATTAAGACGTTGGCGTGGATTAACGCTGGCGGAAAGCG	2100
DB	2041	TATGAAAACGCGCAAGTACGACATTAAGACGTTGGCGTGGATTAACGCTGGCGGAAAGCG	2100


```
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAAAAACAAACACATGAATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAAAAACAAACACATGAATGACAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTAATTAAGGCGGAAACAGCGATGACCATCAAGCGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATTAAGGCGGAAACAGCGATGACCATCAAGCGCCCGTGG 2220
QY 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAATGTTAAACGGTACTCGGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAATGTTAAACGGTACTCGGACCTTC 2280
QY 2281 AAGGTCACCAATCAAAACCGTTCGTTGGCGTGTGAGCGAGGTATTAAGCGCCGCCAGT 2340
Db 2281 AAGGTCACCAATCAAAACCGTTCGTTGGCGTGTGAGCGAGGTATTAAGCGCCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATCAAGGTCG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAACTATCTGCTGACTGATCAAGGTCG 2400
QY 2401 GAAGCGGTAAATAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Db 2401 GAAGCGGTAAATAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
QY 2461 TTGGCGAAAGATCCACGATTTCGCCCCACCATGGAAACCGCCGAGAAAGGTGAAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTTCGCCCCACCATGGAAACCGCCGAGAAAGGTGAAATCATG 2520
QY 2521 CCGAACATCCGCGAGATGTCGCTTCCTGTTGGTATGCGGTGCGTACTCGCGGTGATCAAGCC 2580
Db 2521 CCGAACATCCGCGAGATGTCGCTTCCTGTTGGTATGCGGTGCGTACTCGCGGTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCTGAAAGACGCGCAGACTTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCTGAAAGACGCGCAGACTTAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
QY 2701 GGATCTCTCTCTGTCGCGCCAGCGGCTTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCTCTCTCTGTCGCGCCAGCGGCTTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAAATGTTGTAAGTCTCTTCAGCCAGAGCAGCTCCCTGGTGGCCACAGCTACCCAC 2820
Db 2761 CCGGAAATGTTGTAAGTCTCTTCAGAGCTCTCACTGTTGGCCACAGCTACCCAC 2820
QY 2821 ACGGGTGAAACACCGTATAAATGCCAGAGTGGCGCAATCTTTTAGCCAGAGCAGCAGC 2880
Db 2821 ACGGGTGAAACACCGTATAAATGCCAGAGTGGCGCAATCTTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGGTGGCCATCAACGACCTATCTGCGGAGAGCCATACAACTCTCAGATGTGGC 2940
Db 2881 CTGGTGGCCATCAACGACCTATCTGCGGAGAGCCATACAACTCTCAGATGTGGC 2940
QY 2941 AAGTCTTTCTGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3000
Db 2941 AAGTCTTTCTGAGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3000
QY 3001 CCCTATGTTGTCGGGAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3060
Db 3001 CCCTATGTTGTCGGGAATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3060
QY 3061 CAGGTTACCCACAGCGGTGAAAAACCGTATAAATGCCCAGAGTGGCGCAATCTTTTAGC 3120
Db 3061 CAGGTTACCCACAGCGGTGAAAAACCGTATAAATGCCCAGAGTGGCGCAATCTTTTAGC 3120
QY 3121 CGCAGCGATTAACCTGGTGGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
Db 3121 CAGGCCGGCCACCTGGCGAGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
```

```
QY 3181 CCAGAAATGTGGCAAGTCTTTCTCAACTTCAGGCCATTTGGTTCCTGACCAACGTAATCTAC 3240
Db 3181 CCAGAAATGTGGCAAGTCTTTCTCGGTCTGACAAATCTCGTCCGGCACCAACGTAATCTAC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGGCGGAGTACCCGTACGACGTTCCGGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGGCGGAGTACCCGTACGACGTTCCGGACTACGCT 3300

RESULT 5
AX202428 LOCUS 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 16 from Patent WO0152620.
ACCESSION AX202428
VERSION AX202428.1 GI:15392176
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1 other sequences; artificial sequences.
AUTHORS Barbas, C.P., Stege, J.T., Guan, X. and Dalmia, B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 16 26-JUL-2001;
The Scripps Research Institute (US); SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES
source Location/Qualifiers
1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="Partial sequence of pMal-m3 and zinc finger protein ZFPm3"
```

ORIGIN

```
Query Match 97.2%; Score 3207.2; DB 6; Length 3300;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 CCAGACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATATAGCGCCCGGAGAGA 60
Db 1 CCAGACCATCGAATGGTGCAAAACCTTTCCGGTATGGCATATAGCGCCCGGAGAGA 60
QY 61 GTCAAATTCAGGGTGGTGAATGTGAAACCAAGTATACGATGTGCGAGAGTATGCCG 120
Db 61 GTCAAATTCAGGGTGGTGAATGTGAAACCAAGTATACGATGTGCGAGAGTATGCCG 120
QY 121 GTGTCTCTTATACAGCGTTTCCCGGTGGTGAACCAAGCCAGCCAGCTTTCTGCGAAA 180
Db 121 GTGTCTCTTATACAGCGTTTCCCGGTGGTGAACCAAGCCAGCCAGCTTTCTGCGAAA 180
QY 181 CCGCGGAAAAGTGGAAAGCGCGATGCGGAGCTGAATTTACATTCCTCCACCGCGTGGC 240
Db 181 CCGCGGAAAAGTGGAAAGCGCGATGCGGAGCTGAATTTACATTCCTCCACCGCGTGGC 240
QY 241 AACAAATGCGCGCAACAGTCTGCTGATTTGGCGTTCGCCACTTCAGTCTGGCCCTGC 300
Db 241 AACAAATGCGCGCAACAGTCTGCTGATTTGGCGTTCGCCACTTCAGTCTGGCCCTGC 300
QY 301 ACAGCGCTGCGAAATTTGTCGGCGGATTAATCTCGCGCGATCACTGGGTGCGAGCG 360
Db 301 ACAGCGCTGCGAAATTTGTCGGCGGATTAATCTCGCGCGATCACTGGGTGCGAGCG 360
QY 361 TGGTGTGTGATGTTAGAACGAGCGGCTGAAAGCGGCTGTAAGCGGGTGCACAATC 420
Db 361 TGGTGTGTGATGTTAGAACGAGCGGCTGAAAGCGGCTGTAAGCGGGTGCACAATC 420
QY 421 TTCTCCGCAACCGCTGAGTGGCTGATCAATTAATCTCGCTGGATGACCGAGTGCACA 480
Db 421 TTCTCCGCAACCGCTGAGTGGCTGATCAATTAATCTCGCTGGATGACCGAGTGCACA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTTAATTTCCGCGTATTTCTTGATGTCTGTGACAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTTAATTTCCGCGGTTATTTCTTGATGTCTGTGACAGA 540
```

QY 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
QY 601 TGGTCGCATTGGGTCACAGCAAAATCGCGCTGTTAGCGGCGCCATTAAAGTTCGTCTCGG 660
DB 601 TGGTCGCATTGGGTCACAGCAAAATCGCGCTGTTAGCGGCGCCATTAAAGTTCGTCTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGCTGGCAATAATCTCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGCGTCTGGCTGGCTGGCAATAATCTCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
DB 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 840
DB 781 ATGAGGGCATCGTTCCCACTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 840
QY 841 TCGCGGCCATTACCGAGTCCGGGCTGCGCGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
DB 841 TCGCGGCCATTACCGAGTCCGGGCTGCGCGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCAACCATCAAAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACCAACCATCAAAACAGGATTTTC 960
QY 961 GCCTCTGGGGCAACACGAGTGGAGCCGCTTCTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
DB 961 GCCTCTGGGGCAACACGAGTGGAGCCGCTTCTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTTGCGCGCTCTCACTGGTGAAGAAAAACCACTGGCGGCCAATA 1080
DB 1021 AGGCAATCAGCTGTTGCGCGCTCTCACTGGTGAAGAAAAACCACTGGCGGCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCACTGGCGACGACAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCACTGGCGACGACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
DB 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAATCTCATGTTTGACAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260
DB 1201 GCACAATCTCATGTTTGACAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACGCCATCGGAAGCTGGTATGGCTGTGAGGTCGTAATCACTGCATAAATTCG 1320
DB 1261 TCAGGACGCCATCGGAAGCTGGTATGGCTGTGAGGTCGTAATCACTGCATAAATTCG 1320
QY 1321 TGTCCCTCAAGGCGCACTCCCGTTCTGGATATGTTTTTGGCGCGACATCATACGGTT 1380
DB 1321 TGTCCCTCAAGGCGCACTCCCGTTCTGGATATGTTTTTGGCGCGACATCATACGGTT 1380
QY 1381 CTGGCAATATTTCTGAATGAGCTGTTGACAATTAATCACTCGGCTGTAATGTGTGA 1440
DB 1381 CTGGCAATATTTCTGAATGAGCTGTTGACAATTAATCACTCGGCTGTAATGTGTGA 1440
QY 1441 ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTTAGGTGTTTTCA 1500
DB 1441 ATTGTGAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTTAGGTGTTTTCA 1500
QY 1501 GCACCTTCAACAAAGGACCATAGATTATGAAGAACTGAAGAGGTAACCTGTAATCTGG 1560
DB 1501 GCACCTTCAACAAAGGACCATAGATTATGAAGAACTGAAGAGGTAACCTGTAATCTGG 1560
QY 1561 ATTTAAGCGGATAAAGGCTATACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
DB 1561 ATTTAAGCGGATAAAGGCTATACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620

QY 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGATAAACTGGGAAGAGAAATTTCCACAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACGGTTGAGCATCCGATAAACTGGGAAGAGAAATTTCCACAGGTT 1680
QY 1681 CGCGCAACTGCGGATGGCCCTGACATATCTTCTGGGCACACGCGCTTTGGTGGCTTAC 1740
DB 1681 CGCGCAACTGCGGATGGCCCTGACATATCTTCTGGGCACACGCGCTTTGGTGGCTTAC 1740
QY 1741 GCTCAATCTGCGCTTGGCTGGAATCAACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
DB 1741 GCTCAATCTGCGCTTGGCTGGAATCAACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CGGTTTACCTGGGATGGCTAGCTTACAAACGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
DB 1801 CGGTTTACCTGGGATGGCTAGCTTACAAACGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
QY 1861 GAAGGTTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
DB 1861 GAAGGTTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGTAAGAGCGCGTGAATGTTCAAC 1980
DB 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGTAAGAGCGCGTGAATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTACCTGGCCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
DB 1981 CTGCAAGAACCGTACTTACCTGGCCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGGCGTGGATTAACGCTGGGCGGAAAGCG 2100
DB 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGGCGTGGATTAACGCTGGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAAAAACAAACACATGAAATGCGAGACACCGATTAC 2160
DB 2101 GGTCTGACCTTCTCGTTGACCTGATTAAAAACAAACACATGAAATGCGAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAAGCTGCGCTTTAAAGCGGAAACAGCGATGACCATCAACGGCCGCTGG 2220
DB 2161 TCCATCGCAGAAGCTGCGCTTTAAAGCGGAAACAGCGATGACCATCAACGGCCGCTGG 2220
QY 2221 GCATGGTCAACATCGACACGCAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
DB 2221 GCATGGTCAACATCGACACGCAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGTCACACCATCAACACCGTTTCTGCGCTGCTGAGCGCAGGATTAACCGCCAGT 2340
DB 2281 AAGGTCACACCATCAACACCGTTTCTGCGCTGCTGAGCGCAGGATTAACCGCCAGT 2340
QY 2341 CGAAACAAAGAGCTGGCAAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
DB 2341 CGAAACAAAGAGCTGGCAAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2460
DB 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2460
QY 2461 TTGGCAAGAGTCCACGTTATTCGCCCAACCTGGAAGAAACGCCAGAAAGGTGAAATCATG 2520
DB 2461 TTGGCAAGAGTCCACGTTATTCGCCCAACCTGGAAGAAACGCCAGAAAGGTGAAATCATG 2520
QY 2521 CCGAACATCCCGAGATGTCGCTTTCTGGTATGTCGCTGGTATGTCGCGTATCAACGCC 2580
DB 2521 CCGAACATCCCGAGATGTCGCTTTCTGGTATGTCGCTGGTATGTCGCGTATCAACGCC 2580
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGGAAGACGCGCAGACTAAATTCGAGCTCG 2640
DB 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGGAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAAATAACAAACAAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
DB 2641 AACAAACAAACAAATAACAAACAAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
QY 2701 GGATCCTCTTCTCTGTCGGCCCGAGCGGCCCTCTCGAGCCCGGGGAGAACCCCTATGCTGT 2760

```
Db 2701 GGAATCCTCTCTCTGTGGCCAGCGGCCCTCGAGCCGGGAGAAGCCCTATGCTGT 2760
Qy 2761 CCGGAATGTGGTAAGTCTTTCAGCAGAGCAGCTCCCTGGTGGCCACACAGCGTACCCAC 2820
Db 2761 CCGGAATGTGGTAAGTCTTTCAGCAGATCCTGGCCACCTGGTGGCCACAGCGTACCCAC 2820
Qy 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTTAGCAGAGCAGCAGC 2880
Db 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTTAGCAGCAGCGCTCC 2880
Qy 2881 CTGGTGGCCATCAACGCACTCATCTACTGGCAGAGGAGCATACAAATGTCAGATGTGGC 2940
Db 2881 CTGGTGGCCATCAACGCACTCATCTACTGGCAGAGGAGCATACAAATGTCAGATGTGGC 2940
Qy 2941 AAGTCTTTTCAGTGAATGTGCTGATCTTGGAGGACCAACAGTACTCACACCGGGAGAG 3000
Db 2941 AAGTCTTTTCAGCAGAGCTCCAGCTTGTGGCCACCAAGTACTCACACCGGGAGAG 3000
Qy 3001 CCCTATGCTTGTCCGGAATGTGTAAGTCTTCTCAGAGCTCTCACCTGTGGCCAC 3060
Db 3001 CCCTATGCTTGTCCGGAATGTGTAAGTCTTTCAGCAGAGCAGCTCCCTGTGGCCAC 3060
Qy 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTAGC 3120
Db 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTAGT 3120
Qy 3121 CGCAGCGATAAAGTGGTGGCCATCAACGCACTCATCTGCGAGAGGCCATACAAATGT 3180
Db 3121 GACTGCCGCGACCTTGTGCGCATCAACGCACTCATCTGCGAGAGGCCATACAAATGT 3180
Qy 3181 CCAGATGTGGCAAGTCTTCTCACTTCAGGCCAATTTGGTCCGTCACCAAGTACTCAC 3240
Db 3181 CCAGATGTGGCAAGTCTTCTCCCAATCCAGCCATCTCGTCCGCGACCAACGTAAC 3240
Qy 3241 ACCGGTAAAAAATAGTGGCCAGCGCCAGTACCGTACGAGCTTCGCGACTACGCT 3300
Db 3241 ACCGGTAAAAAATAGTGGCCAGCGCCAGTACCGTACGAGCTTCGCGACTACGCT 3300

RESULT 6
AX284144 7373 bp DNA linear PAT 20-NOV-2001
LOCUS
DEFINITION Sequence 1 from Patent WO0178756.
ACCESSION AX284144
VERSION AX284144.1 GI:17044832
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE
1. Wiederanders, B. and Maubach, G.
AUTHORS Agent for postoperative use after the removal of bone tumours
TITLE Patent: WO 0178756-A 1 25-OCT-2001;
JOURNAL Depuy Biotech Jena GmbH (DE)
FEATURES
Location/Qualifiers
source 1..7373
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="Expressionsvektor pMalc2cchbnp2"
1528..3453
/notes="unnamed protein product; /produkt=Fusionsprotein
MBP Cystatin C BMP-2"
/codon_start=1
/transl_table=11
/protein_id="CAD12160.1"
/db_xref="GI:17044833"
/translation="MKIEGKLVIVINGDKYGNLAVGVKKPFKDTGKIVUHPDKL
EEFPQVAATGDGPDIIIFWADHFGGTAQSGLLAEITPDFAQDKLVPFTWADVRVNG
KLIAYPJAVSALSYIKNDLLPNPKTWEEIPALDKELKAGKSALMFNLQEPYFTWP
LIAADGGYAFKNGYKIDKVDGNAGAKAGLTFILVDLIIKNKHMADTDYSIAEAA
NKGETAMTINGPMAWSNIDTSKVNYGVTLVLPFKGQPSKPFVGLSAGINAAAPNKEL
```

```
misc_feature
/note="Faktor Xa Protease /position=388 (AMINO
ACIDsequenz)"
3067..3108
/misc_feature
/note="/spacer zwischen Cystatin C und BMP-2"
ORIGIN
Query Match 82.0%; Score 2705.8; DB 6; Length 7373;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2719; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATCGTGCAGAAACCTTTCGCGGTATGGCATGATAGCGCCCGAAGAGA 60
Db 1 CCGACACCATCGAATCGTGCAGAAACCTTTCGCGGTATGGCATGATAGCGCCCGAAGAGA 60
Qy 61 GTCMAATCAGGTGGTGAATGTGAACACCACTAAGTTATACGATGTGCGAGAGTATCGCG 120
Db 61 GTCMAATCAGGTGGTGAATGTGAACACCACTAAGTTATACGATGTGCGAGAGTATCGCG 120
Qy 121 GTGTCCTTATCAGACCGTTTCCGCGGTGTGAACAGGCGCAGCAGCTTCTTCGCAAAA 180
Db 121 GTGTCCTTATCAGACCGTTTCCGCGGTGTGAACAGGCGCAGCAGCTTCTTCGCAAAA 180
Qy 181 CGCGGAAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTACATTCCCAACCGGTTGGCAC 240
Db 181 CGCGGAAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTACATTCCCAACCGGTTGGCAC 240
Qy 241 AACCACTGGCGGGGCAACAGTCGTGCTGATTTGGCGTTGCGCACTCCAGTCTGGCCCTGC 300
Db 241 AACCACTGGCGGGGCAACAGTCGTGCTGATTTGGCGTTGCGCACTCCAGTCTGGCCCTGC 300
Qy 301 ACGCGCGTGCAGAAATGTGCGCGGATTAATATCTCGCGCGGATCACTCGGTTGCCAGCG 360
Db 301 ACGCGCGTGCAGAAATGTGCGCGGATTAATATCTCGCGCGGATCACTCGGTTGCCAGCG 360
Qy 361 TGGTGGTGTGATGTAGAAAGCGCGCTCGAAGCTTGAAGCGCGGCGGTGCACAATC 420
Db 361 TGGTGGTGTGATGTAGAAAGCGCGCTCGAAGCTTGAAGCGCGGCGGTGCACAATC 420
Qy 421 TTCTCGCGCAACGGCTCAGTGGGCTGATTAATCACTCCGCTGGATGACCAAGATGCCA 480
Db 421 TTCTCGCGCAACGGCTCAGTGGGCTGATTAATCACTCCGCTGGATGACCAAGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
Db 481 TTGCTGTGGAAGCTGCCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
Qy 541 CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATATTTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTGCATTTGGGTACCGCAAAATCGCTGTTAGCGGGCCATTAAGTCTGCTCTCG 660
Db 601 TGGTGCATTTGGGTACCGCAAAATCGCTGTTAGCGGGCCATTAAGTCTGCTCTCG 660
Qy 661 CGCGTCTGCGTGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAAACGGGAAGCGGAGTGGCATGTCCGGTTTTCAAACCAATGCAAAATGCTGA 780
Db 721 CGGAAACGGGAAGCGGAGTGGCATGTCCGGTTTTCAAACCAATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATCGTTTCCCACTCGCATGCTGTTGCCAAACGATCAGATGGCGCTGGCGCAA 840
Db 781 ATGAGGGCATCGTTTCCCACTCGCATGCTGTTGCCAAACGATCAGATGGCGCTGGCGCAA 840
```

841 TCGCGCCCATTTACCGAGTCCGGCTCGCGTGTGGTGGGATATCTCGGTAGTGGATACG 900
Db |||||
841 TCGCGCCCATTTACCGAGTCCGGCTCGCGTGTGGTGGGATATCTCGGTAGTGGATACG 900
Qy |||||
901 ACGATACCGAAGACAGCTCATGTATATCCCGCGTTTAAACCAACCATCAAAAGGATTTTC 960
Db |||||
901 ACGATACCGAAGACAGCTCATGTATATCCCGCGTTTAAACCAACCATCAAAAGGATTTTC 960
Qy |||||
961 GCTGTGTGGGCAAAACAGCGTGGACCGCTTCTGCAACTCTCTCAGGCCACAGCGGTGA 1020
Db |||||
961 GCTGTGTGGGCAAAACAGCGTGGACCGCTTCTGCAACTCTCTCAGGCCACAGCGGTGA 1020
Qy |||||
1021 AGGCAATCAGCTGTTCGCCCTCTCACTGGTGAAGAAAACCAACCCCTGGCGCCCAATA 1080
Db |||||
1021 AGGCAATCAGCTGTTCGCCCTCTCACTGGTGAAGAAAACCAACCCCTGGCGCCCAATA 1080
Qy |||||
1081 CGCAAAACCGCTCTCCCGCGGTGGCGGATTCATTATGACGCTGGCGACCAAGGTTT 1140
Db |||||
1081 CGCAAAACCGCTCTCCCGCGGTGGCGGATTCATTATGACGCTGGCGACCAAGGTTT 1140
Qy |||||
1141 CCCGACTGGAAGCGGAGTGGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
Db |||||
1201 GCACAATTCATGTTTGACGCTTATCATCGACTGCGAGGTGCACCAATGCTTCTGGCG 1260
Db |||||
1201 GCACAATTCATGTTTGACGCTTATCATCGACTGCGAGGTGCACCAATGCTTCTGGCG 1260
Qy |||||
1261 TCAGGCGGCATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAATCACTGCATTAATTCG 1320
Db |||||
1261 TCAGGCGGCATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAATCACTGCATTAATTCG 1320
Qy |||||
1321 TGTGCTCAAGCGCACCTCCGTTCTGGATAATGTTTTTGGCGCGACATCAACGGTT 1380
Db |||||
1321 TGTGCTCAAGCGCACCTCCGTTCTGGATAATGTTTTTGGCGCGACATCAACGGTT 1380
Qy |||||
1381 CTGGCAATATTTCTGAATAGCTGTTGACATTAATCATCGGCTCGTATATGTGTGA 1440
Db |||||
1381 CTGGCAATATTTCTGAATAGCTGTTGACATTAATCATCGGCTCGTATATGTGTGA 1440
Qy |||||
1441 ATTGTGAGCGGATAACAATTTCCACAGGAACAGCCAGCTCGGTTTAGTGTTTTTCAGGA 1500
Db |||||
1441 ATTGTGAGCGGATAACAATTTCCACAGGAACAGCCAGCTCGGTTTAGTGTTTTTCAGGA 1500
Qy |||||
1501 GCACCTTCAACCAAGGACCATAGATTATGAANAATCGAAGAGGTAACTGGTAATCTGG 1560
Db |||||
1561 ATTACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAAGAT 1620
Db |||||
1561 ATTACCGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAAGAT 1620
Qy |||||
1621 ACCGNAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT 1680
Db |||||
1621 ACCGNAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT 1680
Qy |||||
1681 GCGGCAACTGGCGATGGCCCTGACATTATCTTGGGCAACAGCCGCTTGGTGGCTAC 1740
Db |||||
1681 GCGGCAACTGGCGATGGCCCTGACATTATCTTGGGCAACAGCCGCTTGGTGGCTAC 1740
Qy |||||
1741 GCTCAATCTGGCTGTGGCTGAATATCAACCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
Db |||||
1801 CCGTTTACCTGGATGCCGTAGCTTACACCGCAAGCTGATTCCTTACCCGATCGCTGT 1860
Db |||||
1861 GAAAGCTTATCGTGAATTTTAAACAAAGATCTGTGCGGAAACCGGCCAAAACCTGGGAA 1920
Db |||||
1861 GAAAGCTTATCGTGAATTTTAAACAAAGATCTGTGCGGAAACCGGCCAAAACCTGGGAA 1920

1921 GAGATCCCGCGCTGGATAAAGAACTGAAGAGGTAAGAGCGCGCTGATGTTCAAC 1980
Db |||||
1921 GAGATCCCGCGCTGGATAAAGAACTGAAGAGGTAAGAGCGCGCTGATGTTCAAC 1980
Qy |||||
1981 CTGCAAGAACCGTACTTCCACCTGGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db |||||
1981 CTGCAAGAACCGTACTTCCACCTGGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Qy |||||
2041 TATGAAAAACGGAAGTACGATTAAGAACGTTGGCGCTGGATTAACGCTGGCGGAAAGCG 2100
Db |||||
2041 TATGAAAAACGGAAGTACGATTAAGAACGTTGGCGCTGGATTAACGCTGGCGGAAAGCG 2100
Qy |||||
2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACCAACACATGAATGCGACACCCGATTAC 2160
Db |||||
2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACCAACACATGAATGCGACACCCGATTAC 2160
Qy |||||
2161 TCCATCGCAGAGCTGCCCTTTAATAAAGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db |||||
2161 TCCATCGCAGAGCTGCCCTTTAATAAAGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy |||||
2221 GCATGGTCCAAACATCGACACAGCAAAAGTGAATTAATGTGTAAACGGTACTGCGCACCTTC 2280
Db |||||
2281 AAGGTCACACCATCCAAACCGTTGGTGGCGTCTGAGCGCAGGTATTAACGCGCCGAGT 2340
Db |||||
2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACATGAAGGTTCTG 2400
Db |||||
2401 GAAAGCGTTAATAAAGACAAAACCGCTGGTGGCTGAGCTGAACTCTTACGAGAAAGAG 2460
Db |||||
2461 TTGGGAAAGATCCACGATTTGCCCGCACCATGGAAGACGCCAGAAAGGTGAATCATG 2520
Db |||||
2461 TTGGGAAAGATCCACGATTTGCCCGCACCATGGAAGACGCCAGAAAGGTGAATCATG 2520
Qy |||||
2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCCGTGCTGCTGCTGATCAACGCC 2580
Db |||||
2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCCGTGCTGCTGCTGATCAACGCC 2580
Qy |||||
2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG 2640
Db |||||
2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG 2640
Qy |||||
2641 AACAAACAACAATAACAATAACAACAACTCGGGATCGAGGGAAGGATTTTCAAGATTC 2700
Db |||||
2641 AACAAACAACAATAACAATAACAACAACTCGGGATCGAGGGAAGGATTTTCAAGATTC 2700
Qy |||||
2701 GGATCCTCTTCTCTGTGGCGCCAGCGCCCTTCGAGGCCGG 2741
Db |||||
2701 GGATCCTCAGTCCCGGCAAGCCCGCCCTGTTGGGAGG 2741

RESULT 7
AXI72306
LOCUS AXI72306 8101 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 5 from Patent WO014467.
ACCESSION AXI72306
VERSION AXI72306.1 GI:14597487
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Goubin-Gramatica,F., Ducommun,B. and Prevost,G.
TITLE Method for obtaining human cdc25 phosphatases and method for
identifying human cdc25 phosphatase modulators
JOURNAL Patent: WO 014467-A 5 21-JUN-2001;
SOCIETE DE CONSEILS DE RECHERCHES ET D'APPLICATIONS SCIENTIFIQUES

FEATURES		ORIGIN	Query Match Best Local Similarity Matches 2707; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
source	Location/Qualifiers		
1. .8101	/organism="synthetic construct"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:32630"		
	/note="souche JM109 / pMAL - Hs Cdc25C"		
QY	1	CGACACCATCGAATGGTGC AAAACCTTTTCGGGTATGGCATGATGAGCCCGGAAGAGA	60
Db	1	CGACACCATCGAATGGTGC AAAACCTTTTCGGGTATGGCATGATGAGCCCGGAAGAGA	60
QY	61	GTCAATTCAGGGTGGTGAATGTGAACCAAGTAACTTATACGATGTCAGAGTATGCCG	120
Db	61	GTCAATTCAGGGTGGTGAATGTGAACCAAGTAACTTATACGATGTCAGAGTATGCCG	120
QY	121	GTGTCCTTATCAGACCGCTTTCCCGGTGGTGAACCGGCAGCCACGCTTTCTGCGAAA	180
Db	121	GTGTCCTTATCAGACCGCTTTCCCGGTGGTGAACCGGCAGCCACGCTTTCTGCGAAA	180
QY	181	CGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTACATTCCTCCAAACCGGTGGC	240
Db	181	CGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTACATTCCTCCAAACCGGTGGC	240
QY	241	AACAACTCGCGGCGAAACAGTCGTTGCTGATGGGTTGCCACTCCAGTCTGGCCCTGC	300
Db	241	AACAACTCGCGGCGAAACAGTCGTTGCTGATGGGTTGCCACTCCAGTCTGGCCCTGC	300
QY	301	ACGCGCGCTCGCAATTTCTCGCGCGATTAATCTCGCGCGATCAACTCGGCTGCCAGC	360
Db	301	ACGCGCGCTCGCAATTTCTCGCGCGATTAATCTCGCGCGATCAACTCGGCTGCCAGC	360
QY	361	TGGTGGTTCGATGTGAACGAGCGCGTTCGAGCGCTGTAAGCGCGGTGCACAATC	420
Db	361	TGGTGGTTCGATGTGAACGAGCGCGTTCGAGCGCTGTAAGCGCGGTGCACAATC	420
QY	421	TTCTCGCCCAACGGTCAAGTGGCTGATCAATTAATCTCGCGCGATCAACTCGGCTGCC	480
Db	421	TTCTCGCCCAACGGTCAAGTGGCTGATCAATTAATCTCGCGCGATCAACTCGGCTGCC	480
QY	481	TGCTGTGGAAAGCTGCCTGCATTAATGTTCGGCGTTATTTCTTGAATCTCTGACCC	540
Db	481	TGCTGTGGAAAGCTGCCTGCATTAATGTTCGGCGTTATTTCTTGAATCTCTGACCC	540
QY	541	CACCATCAACAGTATTATTTCTCCCATGAAGCGGTACCGCATCGGCGGTGGAGCATC	600
Db	541	CACCATCAACAGTATTATTTCTCCCATGAAGCGGTACCGCATCGGCGGTGGAGCATC	600
QY	601	TGGTGCATTTGGGTACACAGCAATCGCGCTGTTAGCGGGCCATTAAGTTCTGTCGG	660
Db	601	TGGTGCATTTGGGTACACAGCAATCGCGCTGTTAGCGGGCCATTAAGTTCTGTCGG	660
QY	661	CGGCTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATG	720
Db	661	CGGCTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATG	720
QY	721	CGGAACGGGAGCGACTGGAGTGCATGTCGGTTTTCACAAACCATGCAATGTCTGA	780
Db	721	CGGAACGGGAGCGACTGGAGTGCATGTCGGTTTTCACAAACCATGCAATGTCTGA	780
QY	781	ATGAGGGCATCGTTCCCACTCGCATGCTGTTGCAACGATCAGATGGCGCTGGCGCAA	840
Db	781	ATGAGGGCATCGTTCCCACTCGCATGCTGTTGCAACGATCAGATGGCGCTGGCGCAA	840
QY	841	TGCGCGCATTCACGAGTCCGGGCTGGCGGTGGTGGGATATCTCGTAGTGGGATACG	900
Db	841	TGCGCGCATTCACGAGTCCGGGCTGGCGGTGGTGGGATATCTCGTAGTGGGATACG	900

QY	901	ACGATACCGAAGACAGCTCATGTTATATATCCGCCGTTAAACCAACCATCAAAACAGGATTTTC	960
Db	901	ACGATACCGAAGACAGCTCATGTTATATATCCGCCGTTAAACCAACCATCAAAACAGGATTTTC	960
QY	961	GCCTGCTGGGGCAAAACCAAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
Db	961	GCCTGCTGGGGCAAAACCAAGCGTGGACCGCTTGTCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
QY	1021	AGGCAATCAGCTGTTGCCGCTCTCACTGTGTAAGAAAAAACCAACCTGGCGGCCCAATA	1080
Db	1021	AGGCAATCAGCTGTTGCCGCTCTCACTGTGTAAGAAAAAACCAACCTGGCGGCCCAATA	1080
QY	1081	CGCAAAACCGCTCTCCCGCGCGTGGCCGATTCAATTAATGACGCTGGCAGCAGAGTTT	1140
Db	1081	CGCAAAACCGCTCTCCCGCGCGTGGCCGATTCAATTAATGACGCTGGCAGCAGAGTTT	1140
QY	1141	CCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG	1200
Db	1141	CCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG	1200
QY	1201	GCACAAATTCATGTTTGGACAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG	1260
Db	1201	GCACAAATTCATGTTTGGACAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGCG	1260
QY	1261	TCAGGACAGCCATCGAAAGCTGTGATGCTGTGCAAGTGTAAATCACTGCATAATTCG	1320
Db	1261	TCAGGACAGCCATCGAAAGCTGTGATGCTGTGCAAGTGTAAATCACTGCATAATTCG	1320
QY	1321	TGTCGCTCAAGGGCGCATCCCGTTCTGGATAATGTTTTTGGCGCGACATCATTAACGGTT	1380
Db	1321	TGTCGCTCAAGGGCGCATCCCGTTCTGGATAATGTTTTTGGCGCGACATCATTAACGGTT	1380
QY	1381	CTGCAAAATTTCTGAAATGAGCTGTGTAACAATTAATCATCGGCTGTAATGTTGTTGGA	1440
Db	1381	CTGCAAAATTTCTGAAATGAGCTGTGTAACAATTAATCATCGGCTGTAATGTTGTTGGA	1440
QY	1441	ATTGTGAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA	1500
Db	1441	ATTGTGAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA	1500
QY	1501	GCATTCACCAACAGGACCATGATTTATGAAAATCGAAGAGGTAACTGGTAAATCTGG	1560
Db	1501	GCATTCACCAACAGGACCATGATTTATGAAAATCGAAGAGGTAACTGGTAAATCTGG	1560
QY	1561	ATTAAACGGCGATAAAGCTATAACGGTCTCGCTGAGTTCGTAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAAACGGCGATAAAGCTATAACGGTCTCGCTGAGTTCGTAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCGAAGAGAAATTCACACAGGTT	1680
Db	1621	ACCGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCGAAGAGAAATTCACACAGGTT	1680
QY	1681	CGGCAACTGGCGATGGCGCTGACATTTCTCTGGGCAACAGCCGCTTTGGTGGCTAC	1740
Db	1681	CGGCAACTGGCGATGGCGCTGACATTTCTCTGGGCAACAGCCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCTGTTGGCTGAAATCACCCCGCAAAAGCGTTCCAGGACAAAGCTGAT	1800
Db	1741	GCTCAATCTGGCTGTTGGCTGAAATCACCCCGCAAAAGCGTTCCAGGACAAAGCTGAT	1800
QY	1801	CCGTTTACCTGGGATCCGTACGTTACACCGCAAGCTGATTGCTTACCCGATCGCTGTT	1860
Db	1801	CCGTTTACCTGGGATCCGTACGTTACACCGCAAGCTGATTGCTTACCCGATCGCTGTT	1860
QY	1861	GAAGCGTTATCGCTGATTTTAAACAGATCTGCTCCGCAACCCCGCAAAACCTGGGAA	1920
Db	1861	GAAGCGTTATCGCTGATTTTAAACAGATCTGCTCCGCAACCCCGCAAAACCTGGGAA	1920
QY	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGGAAAGGTAAAGAGCGCTGATGTTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGGAAAGGTAAAGAGCGCTGATGTTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTTCACCTGGCGCTGATTTGCTGCTGACCGGGGTTATGCTTCAAG	2040

Db	1981	CTGCAAGAACCGTACTTACCTGCCCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG	2040
Qy	2041	TATGAAAACGGCAAGTACGACATTTAAAGACGCTGGCGTGGATACGCTGGCGGCAAAAGCG	2100
Db	2041	TATGAAAACGGCAAGTACGACATTTAAAGACGCTGGCGTGGATACGCTGGCGGCAAAAGCG	2100
Qy	2101	GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAAANAACAACATGAATGAGACACCGGATTAC	2160
Db	2101	GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAAANAACAACATGAATGAGACACCGGATTAC	2160
Qy	2161	TCCATCGCAAGCTGCTTTTAAAGAGCGAAACAGCGATGACCATCAACGGCCCGTGG	2220
Db	2161	TCCATCGCAAGCTGCTTTTAAAGAGCGAAACAGCGATGACCATCAACGGCCCGTGG	2220
Qy	2221	GCATGGTCCAAATCGACACCAAGTGAATTTATGGTGAACGGTACTGCGGACCTTC	2280
Db	2221	GCATGGTCCAAATCGACACCAAGTGAATTTATGGTGAACGGTACTGCGGACCTTC	2280
Qy	2281	AAGGGTCAACCATCCAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAACGGCCCGT	2340
Db	2281	AAGGGTCAACCATCCAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAACGGCCCGT	2340
Qy	2341	CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG	2400
Db	2341	CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG	2400
Qy	2401	GAAAGCGGTTAATAAGACAAACCGCTGGGTGGCGTAGCGCTGAAGTCTTACGAGGAAG	2460
Db	2401	GAAAGCGGTTAATAAGACAAACCGCTGGGTGGCGTAGCGCTGAAGTCTTACGAGGAAG	2460
Qy	2461	TTGGCGAAAGATCCACGTTATTCGCGCCACCATGGAAAAACCCCAAGAGGTGAATCATG	2520
Db	2461	TTGGCGAAAGATCCACGTTATTCGCGCCACCATGGAAAAACCCCAAGAGGTGAATCATG	2520
Qy	2521	CCGAACATCCCGACAGATGTCGGCTTCTGGTATGCCGTGGGTACTCGGTGATCAACGCC	2580
Db	2521	CCGAACATCCCGACAGATGTCGGCTTCTGGTATGCCGTGGGTACTCGGTGATCAACGCC	2580
Qy	2581	GCCAGCGGTGCTGAGACTGTCGATGAAGCCCTGAAAGACGCGACGACTAATTCGAGCTCG	2640
Db	2581	GCCAGCGGTGCTGAGACTGTCGATGAAGCCCTGAAAGACGCGACGACTAATTCGAGCTCG	2640
Qy	2641	AACAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC	2700
Db	2641	AACAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTCAGAAATTC	2700
Qy	2701	GGATCCTCT	2709
Db	2701	GGATCCTCT	2709
RESULT 8			
AX377531			
LOCUS	AX377531	6648 bp	DNA linear PAT 18-MAR-2002
DEFINITION	Sequence 8 from Patent WO0212553.		
ACCESSION	AX377531		
VERSION	AX377531.1 GI:19573717		
KEYWORDS	synthetic construct		
SOURCE	synthetic construct		
ORGANISM	other sequences; artificial sequences.		
REFERENCE	1		
AUTHORS	Kappel,A., Polakowski,T., Pignot,M., Windhab,N., Behrens,dorf,H. and Muth,J.		
TITLE	Method for detecting mutations in nucleotide sequences		
JOURNAL	Patent: WO 0212553-A 8 14-FEB-2002;		
FEATURES	Nanogen Recognomics GmbH (DE)		
	Location/Qualifiers		
	1..6648		
	/organism="synthetic construct"		
source			
/mol_type="unassigned DNA"			
/db_xref="taxon:32630"			

/note="Plasmid pMAL-c2x"

ORIGIN	Query Match	81.8%	Score 2701;	DB 6;	Length 6648;
	Best Local Similarity	99.8%	Pred. No. 0;		
	Matches 2704;	Conservative 0;	Mismatches 5;	Indels 0;	Gaps 0;
Qy	1	CGACACCATCGAATGGTGTGCAAAACCTTTTCGGGGTATGGCATGATGAGCCCGGAAAGAGA	60		
Db	1	CGACACCATCGAATGGTGTGCAAAACCTTTTCGGGGTATGGCATGATGAGCCCGGAAAGAGA	60		
Qy	61	GTCRAATTCAGGGTGGTGAATGTGAAACCAAGTAACTTATACGATCTCCAGAGTATGCGG	120		
Db	61	GTCRAATTCAGGGTGGTGAATGTGAAACCAAGTAACTTATACGATCTCCAGAGTATGCGG	120		
Qy	121	GTGTCCTCTTATCAGACCCGTTTCCCGCGTGGTGAACCAAGCCAGCCACGTTTCTGCGAATA	180		
Db	121	GTGTCCTCTTATCAGACCCGTTTCCCGCGTGGTGAACCAAGCCAGCCACGTTTCTGCGAATA	180		
Qy	181	CGCGGAAAAAGTGAAGCGGGATGGCGGAGCTGAAATTAATTCCTCAACCGCGTGGCAC	240		
Db	181	CGCGGAAAAAGTGAAGCGGGATGGCGGAGCTGAAATTAATTCCTCAACCGCGTGGCAC	240		
Qy	241	AACAACTCGCGGGCAACAGTCGTTGCTGATTTGGGCTTGCACCTCCAGTCTGGCCCTGC	300		
Db	241	AACAACTCGCGGGCAACAGTCGTTGCTGATTTGGGCTTGCACCTCCAGTCTGGCCCTGC	300		
Qy	301	ACGCGCCCTCGCAATTTGTCGCGCGATTAATCTCGCGCGCATCAACTGGGTGGCAGCG	360		
Db	301	ACGCGCCCTCGCAATTTGTCGCGCGATTAATCTCGCGCGCATCAACTGGGTGGCAGCG	360		
Qy	361	TGGTGGTGTGATGTAAGAACGAGCGGCTGAAAGCCTGTAAAGCGGCGTGCACAAATC	420		
Db	361	TGGTGGTGTGATGTAAGAACGAGCGGCTGAAAGCCTGTAAAGCGGCGTGCACAAATC	420		
Qy	421	TTCTCGCCCAACGCTGAGTGGGCTGATTAATCTATCCGCTGATGACAGGATGCCA	480		
Db	421	TTCTCGCCCAACGCTGAGTGGGCTGATTAATCTATCCGCTGATGACAGGATGCCA	480		
Qy	481	TTGCTGTGGAAGCTCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCAG	540		
Db	481	TTGCTGTGGAAGCTCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCAG	540		
Qy	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600		
Db	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600		
Qy	601	TGGTGGCATTTGGTTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTTCTCG	660		
Db	601	TGGTGGCATTTGGTTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTTCTCG	660		
Qy	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720		
Db	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720		
Qy	721	CGGAAACGGGAAGCGACATGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA	780		
Db	721	CGGAAACGGGAAGCGACATGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA	780		
Qy	781	ATGAGGGCATCGTTCCCACTGGGATGCTGGTTCGCAACGATCAGATGGCGCTGGGGCGAA	840		
Db	781	ATGAGGGCATCGTTCCCACTGGGATGCTGGTTCGCAACGATCAGATGGCGCTGGGGCGAA	840		
Qy	841	TGCGCGCATTAACCGAGTCCCGGCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	900		
Db	841	TGCGCGCATTAACCGAGTCCCGGCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	900		
Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATGCAAAAGGATTTTC	960		
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACCAACCATGCAAAAGGATTTTC	960		
Qy	961	GGCTGCTGGGGCAACCAAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCAGGCGGTGA	1020		
Db	961	GGCTGCTGGGGCAACCAAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCAGGCGGTGA	1020		

Db 961 GCCTGCTGGGCAAAACCGAGCTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGTGA 1020
Qy 1021 AGGCAATCAGCTGTGGCCCGTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGGCCCGTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGGCTGGCCGATTCATTAATGACGTGGCAGACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGCTGGCCGATTCATTAATGACGTGGCAGACAGGTTT 1140
Qy 1141 CCGACTGGAAGCGGCGAGTCAGCGCAACCAATTAATGTGAGTTAGTCACTCAATTAG 1200
Db 1141 CCGACTGGAAGCGGCGAGTCAGCGCAACCAATTAATGTGAGTTAGTCACTCAATTAG 1200
Qy 1201 GCACAATTTCTATGTTGACAGCTTATCATGACCTGCACGGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTTCTATGTTGACAGCTTATCATGACCTGCACGGTGCACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGACGCATCGGAAGCTGGTATGCTGTGAGTGCAGTTCGTAATCACTGCAATTTTCG 1320
Db 1261 TCAGGACGCATCGGAAGCTGGTATGCTGTGAGTGCAGTTCGTAATCACTGCAATTTTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTCTGATATGTTTTCGCGGACATCATACGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTGATATGTTTTCGCGGACATCATACGTT 1380
Qy 1381 CTGGCAATTTCTGAATGAGCTGTTGACATTAATCATCGGCTCGTAAATCACTGCAATTTTCG 1440
Db 1381 CTGGCAATTTCTGAATGAGCTGTTGACATTAATCATCGGCTCGTAAATCACTGCAATTTTCG 1440
Qy 1441 ATTGTAGCGGATAACAATTTACACAGCAACGACGCTGGTTAGGTTTTCACGA 1500
Db 1441 ATTGTAGCGGATAACAATTTACACAGCAACGACGCTGGTTAGGTTTTCACGA 1500
Qy 1501 GCACCTTCAACCAAGGACCATAGATTAATGAAGCTGAAGAGTAACTGCTAATCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTAATGAAGCTGAAGAGTAACTGCTAATCTGG 1560
Qy 1561 ATTAACGGCGATAAGGCTATAACGCTCTCGTGAAGTCGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAGGCTATAACGCTCTCGTGAAGTCGTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGACATCCGATTAACCTGGAAGAGAAATTCGACAGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGACATCCGATTAACCTGGAAGAGAAATTCGACAGTT 1680
Qy 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCTGTGGCTGAATATCACCCGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTGGCTGAATATCACCCGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGATGCGTACGTTTACACCGCAAGCTGATGCTTACCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGTACGTTTACACCGCAAGCTGATGCTTACCGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCGCAACCCGCAAAACCTCGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCGCAACCCGCAAAACCTCGGAA 1920
Qy 1921 GAGATCCGCGCTGGATAAGAACTGAAGCGAAAGTGAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGCGCTGGATAAGAACTGAAGCGAAAGTGAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGACGGGGTTATGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGCTGACGGGGTTATGGTTCAAG 2040
Qy 2041 TATGAAAACCGCAAGTACGACATTAAAGACGTGGGCTGGATACGCTGCGCGAAAGCG 2100
Db 2041 TATGAAAACCGCAAGTACGACATTAAAGACGTGGGCTGGATACGCTGCGCGAAAGCG 2100

Qy 2101 GGTCTGACCTTCCTGTTGACCTGATTTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCCTGTTGACCTGATTTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAAGCTGCTTTTAAATAAGGCGAAACAGCATGACATCAACGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAAGCTGCTTTTAAATAAGGCGAAACAGCATGACATCAACGCGCCGCTGG 2220
Qy 2221 GCATGGTCCAAACATCGACACACAGCAAAAGTGAATTTATGGTGAACGCTTACGCGACCTTC 2280
Db 2221 GCATGGTCCAAACATCGACACACAGCAAAAGTGAATTTATGGTGAACGCTTACGCGACCTTC 2280
Qy 2281 AAGGCTCAACCATCAACACGTTTGGTGGCGTGTGAGCGCAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGCTCAACCATCAACACGTTTGGTGGCGTGTGAGCGCAGGTATTAACGCGCCAGT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGCTCG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGCTCG 2400
Qy 2401 GAAGCGTTTAAATAAGCAAAACCGCTGGTGGCTAGCGTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGTTTAAATAAGCAAAACCGCTGGTGGCTAGCGTGAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGGCGAAAGATCCAGCTATTTGCCGCCACCATGMAAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGCTATTTGCCGCCACCATGMAAACCGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGAGATGTCCGCTTTCTGGTATGCCGTGCTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGAGATGTCCGCTTTCTGGTATGCCGTGCTACTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGCTCGTCAGACTGTCGATGAAGCCTGAAGACCGCGAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGCTCGTCAGACTGTCGATGAAGCCTGAAGACCGCGAGACTAATTCGAGCTCG 2640
Qy 2641 AACAAACAAACAAATAACAAATAACAACTCGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACAAACAAACAAATAACAAATAACAACTCGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Qy 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709

RESULT 9
A91965 LOCUS A91965 7475 bp DNA circular PAT 22-JAN-2000
DEFINITION Sequence 1 from Patent WO9822496.
ACCESSION A91965
VERSION A91965.1 GI:6740811
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 7475)
AUTHORS Attwood,M.R. and Hurst,D.N.
TITLE ANTIVIRAL PEPTIDE DERIVATIVES
JOURNAL Patent: WO 9822496-A 1 28-MAY-1998;
HOFFMANN LA ROCHE (CH)
FEATURES
source Location/Qualifiers
1..7475
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN
Query Match 81.8%; Score 2700; DB 6; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGATCGTGCATAAACCTTTCCGGGTATGCGGTATGCGATGATAGCGCCCGGAAGAGA 60

|||||
1 CGGACACCATCGAATGGTGC AAAACCTTTCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
61 GTCAATTCAGGGTGGTGAATGTGAACACCAAGTAAACGTTATACGATGTCGCAGAGTATGCCG 120
61 GTCAATTCAGGGTGGTGAATGTGAACACCAAGTAAACGTTATACGATGTCGCAGAGTATGCCG 120
121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACCAAGCCAGCCACGTTTCTGGGAAAA 180
121 GTGTCTCTTATCAGACCGTTTCCCGGTGGTGAACCAAGCCAGCCACGTTTCTGGGAAAA 180
181 CSCGGGAAAAAGTGAAGCGGCGATGGGGAGCTGAATTTACATTCCTCAACCCGCTGGGCAC 240
181 CSCGGGAAAAAGTGAAGCGGCGATGGGGAGCTGAATTTACATTCCTCAACCCGCTGGGCAC 240
241 AACAACTGCGCGGCAAAACAGTGGTGTCTGATTTGGGGTTGGCCACCTCCAGTCTGGGCCCTGC 300
241 AACAACTGCGCGGCAAAACAGTGGTGTCTGATTTGGGGTTGGCCACCTCCAGTCTGGGCCCTGC 300
301 AGCGCGGTGCGCAAAATTTGTCGGCGGATTAATCTTCGCGCCGATCAACTGGGTGCGACGG 360
301 AGCGCGGTGCGCAAAATTTGTCGGCGGATTAATCTTCGCGCCGATCAACTGGGTGCGACGG 360
361 TGGTGGTGTGATGTAGAACCAAGCGGCTCGAAGCCTGTAAAGCGGCGGTGCAACAATC 420
361 TGGTGGTGTGATGTAGAACCAAGCGGCTCGAAGCCTGTAAAGCGGCGGTGCAACAATC 420
421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTTAACTATCCGCTGGATGACCAAGGATGCCA 480
421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTTAACTATCCGCTGGATGACCAAGGATGCCA 480
481 TTGCTGTGGAAGCTGCTGCACTAATGTTCGCGGCTTATTTCTGATGCTCTGACCCAGA 540
481 TTGCTGTGGAAGCTGCTGCACTAATGTTCGCGGCTTATTTCTGATGCTCTGACCCAGA 540
541 CACCCATCAACAGTATTTTCTCCATGAAAGCGGTACGCGACTGGGCGTGGAGCATC 600
541 CACCCATCAACAGTATTTTCTCCATGAAAGCGGTACGCGACTGGGCGTGGAGCATC 600
601 TGGTGCATTTGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTCTGTCTCGG 660
601 TGGTGCATTTGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTCTGTCTCGG 660
661 CGCGTCTGCGTCTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
661 CGCGTCTGCGTCTGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
721 CGGAACGGGAAGCGCACTGGAGTGCATGTCGGTTCGGTTTCAACAAACCATGCAAAATGCTGA 780
721 CGGAACGGGAAGCGCACTGGAGTGCATGTCGGTTCGGTTTCAACAAACCATGCAAAATGCTGA 780
781 ATGAGGGCATCGTCCCACTGCGATGCTGGTTGCCAAACGATCAGATGGGCTGGGCGCAA 840
781 ATGAGGGCATCGTCCCACTGCGATGCTGGTTGCCAAACGATCAGATGGGCTGGGCGCAA 840
841 TCGCGGCCATTACCGAGTCCGGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATAG 900
841 TCGCGGCCATTACCGAGTCCGGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATAG 900
901 AGGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAATCAAAAGGATTTTC 960
901 AGGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAATCAAAAGGATTTTC 960
901 AGGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAATCAAAAGGATTTTC 960
901 AGGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAATCAAAAGGATTTTC 960
961 GCCTGCTGGGGCAAAACACGCGTGGACCGCTGCTGCAACTCTCTCAGGSCCAGGCGGTGA 1020
961 GCCTGCTGGGGCAAAACACGCGTGGACCGCTGCTGCAACTCTCTCAGGSCCAGGCGGTGA 1020
1021 AGGGCAATCAGCTGTGGCCGCTCACTGGTGAAGAAAAACCAACCTTGGCGCCCAATA 1080
1021 AGGGCAATCAGCTGTGGCCGCTCACTGGTGAAGAAAAACCAACCTTGGCGCCCAATA 1080
1081 CGCAACCGCCCTCTCCCGCGGCTTGGCGGATTCATTAATGAGCTGGGCAACAGAGGTTT 1140
|||||

Db 1081 CGCAACCGCCCTCTCCCGCGCGTGGCCGATTCATTAATGAGCTGGCAGCAGAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAAATTCATGTTTGACAGCTTATCATCGACTGCGGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATTCATGTTTGACAGCTTATCATCGACTGCGGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACGCCATCGGAAGCTGTGGTGTGTCAGGTGCGTAAATCACTGCGATAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTGTGTCAGGTGCGTAAATCACTGCGATAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACCTCCCGTTCGTGATAATGTTTTTGGCGCGACATCAATACGTT 1380
Db 1321 TGTGCTCAAGCGGCACCTCCCGTTCGTGATAATGTTTTTGGCGCGACATCAATACGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGTGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGTGA 1440
QY 1441 ATTGTGAGCGGATAACAAATTTTCAACAGGAACAGCCAGTCCGTTTAGTGTTTTTCA 1500
Db 1441 ATTGTGAGCGGATAACAAATTTTCAACAGGAACAGCCAGTCCGTTTAGTGTTTTTCA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGGTAAATCTGGTAATCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGGTAAATCTGGTAATCTGG 1560
QY 1561 ATTTAAACGGCGATTAAGGCTATAACCGTCTCGGTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTTAAACGGCGATTAAGGCTATAACCGTCTCGGTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTTAAAGTCACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCGCCAGGTT 1680
Db 1621 ACCGGAATTTAAAGTCACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCGCCAGGTT 1680
QY 1681 GCGGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGTGGCTGAATAATCAACCCGGGCAAGAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGTGGCTGAATAATCAACCCGGGCAAGAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTGCTTACCCGATCCGCTTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTGCTTACCCGATCCGCTTT 1860
QY 1861 GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGACGCGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGACGCGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGACGTTGGCGCTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGACGTTGGCGCTGGATTAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGGTTGACCTGATTAATAACAAACACATGAATGCAAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGGTTGACCTGATTAATAACAAACACATGAATGCAAGACACCGATTAC 2160
QY 2161 TCCATCGGAGAAAGCTGCGCTTTTAAAGCGGNAACAGGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGGAGAAAGCTGCGCTTTTAAAGCGGNAACAGGATGACCATCAACGGCCCGTGG 2220

```
QY 2221 GCATGGTCCACATCGACACCGCAAGTGAATTATGGTGTAAACGCTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACCGCAAGTGAATTATGGTGTAAACGCTACTGCCGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAAAGCCGCCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAAAGCCGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACCTATCTGCTCACTGATGAGGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACCTATCTGCTCACTGATGAGGTCTG 2400
QY 2401 GAAGCGGTAAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTAAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGCTATTGCGCGCCACCATGGAAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGCTATTGCGCGCCACCATGGAAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTGTGACATGTCGATGAAGCCCTTGAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTGTGACATGTCGATGAAGCCCTTGAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACACACACACATAACAAATAACAAACCTCGGATCGAGGAGGATTCAGAAATTC 2700
Db 2641 AACACACACACATAACAAATAACAAACCTCGGATCGAGGAGGATTCAGAAATTC 2700

RESULT 10
AR031992
LOCUS AR031992 7475 bp DNA linear PAT 29-SEP-1999
DEFINITION Sequence 1 from patent US 5866684.
ACCESSION AR031992
VERSION AR031992.1 GI:5946281
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 7475)
AUTHORS Attwood,M.Richard., Hurst,D.Nigel., Jones,P.Stephen.,
Kay,P.Brittain., Raynham,T.Michael. and Wilson,P.Xavier.
TITLE Peptidyl inhibitors of viral proteases
JOURNAL Patent: US 5866684-A 1 02-FEB-1999;
FEATURES
source
1..7475
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match 81.8%; Score 2700; DB 6; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGTCAAAACCTTTCGCGTATGGCATGATAGGCCCGGAGAGA 60
Db 1 CCGACACCATCGAATGGTGTCAAAACCTTTCGCGTATGGCATGATAGGCCCGGAGAGA 60
QY 61 GTCAATTCAGGGTGGTGAATGTGAAACAGTAAGCTTATACGATGTGCGAGAGTATGCCG 120
Db 61 GTCAATTCAGGGTGGTGAATGTGAAACAGTAAGCTTATACGATGTGCGAGAGTATGCCG 120
QY 121 GTGTCTCTTATCACACGTTTCCCGGTGGTGAACAGGCCAGCCAGTTTCTGCGAAAA 180
Db 121 GTGTCTCTTATCACACGTTTCCCGGTGGTGAACAGGCCAGCCAGTTTCTGCGAAAA 180
QY 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTATCTCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTATCTCCCAACCGCGTGGCAC 240
```

```
Db 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTATCTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAAAACAGTCTGTTGCTGATTTGGCGTTGCAACCTCTCAGTCTGCCCTTC 300
Db 241 AACAACTGGCGGGCAAAACAGTCTGTTGCTGATTTGGCGTTGCAACCTCTCAGTCTGCCCTTC 300
QY 301 ACGCCCGTTCGCAAAATTTGTCGCGCGATTAATCTCTCGCGCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCCCGTTCGCAAAATTTGTCGCGCGATTAATCTCTCGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGTGTGTTCGATGGTGAAGACGAAAGCGCGCTCGAAGCTCTGTAAGCGCGCTGCACAATC 420
Db 361 TGTGTGTTCGATGGTGAAGACGAAAGCGCGCTCGAAGCTCTGTAAGCGCGCTGCACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACACAGGATCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACACAGGATCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACCTAATCTTCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACCTAATCTTCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
QY 541 CACCCATCAACAGATTAATTTCTCCATGAAGACGGTACGCGACTGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGATTAATTTCTCCATGAAGACGGTACGCGACTGGCGTGGAGCATC 600
QY 601 TGGTCGCAATTTGGGTCAACAGCAATCGCGCTGTTAGCGGGCCCATTAAGTCTGTCTCGG 660
Db 601 TGGTCGCAATTTGGGTCAACAGCAATCGCGCTGTTAGCGGGCCCATTAAGTCTGTCTCGG 660
QY 661 CGGCTCTGCGCTCTGGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCGATG 720
Db 661 CGGCTCTGCGCTCTGGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCGATG 720
QY 721 CGGAAACGGGAAGCGCACTGGAATGCCATGTCGGTTTTCACAAACCATCAAAATGCTGA 780
Db 721 CGGAAACGGGAAGCGCACTGGAATGCCATGTCGGTTTTCACAAACCATCAAAATGCTGA 780
QY 781 ATGAGGCGATCGTTCCCACTCGCATGCTGTTGCCAAGATCAGATGGCGCTGGGCGCAA 840
Db 781 ATGAGGCGATCGTTCCCACTCGCATGCTGTTGCCAAGATCAGATGGCGCTGGGCGCAA 840
QY 841 TGGCGCGCATTAACGAGTCCGGCTCGCGCTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGCGCATTAACGAGTCCGGCTCGCGCTTGGTGGGATATCTCGGTAGTGGGATACG 900
QY 901 ACATACCGGAAGACAGCTCATGTTATATCCGCGCTTAACCAACATCAACAGGATTTTC 960
Db 901 ACATACCGGAAGACAGCTCATGTTATATCCGCGCTTAACCAACATCAACAGGATTTTC 960
QY 961 GCCTGTGGGGCAAAACCGAGCTGGACCGCTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGTGGGGCAAAACCGAGCTGGACCGCTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
QY 1021 AGGCAATCAGTGTGTTGCCGCTCTCACTGGTGAAGAAAAAACCCCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGTGTGTTGCCGCTCTCACTGGTGAAGAAAAAACCCCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTGGCGGATTCATTAATGAGCTGGACGACGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTGGCGGATTCATTAATGAGCTGGACGACGAGTTT 1140
QY 1141 CCCGACTGGAAGAGCGGCGAGTGAAGCAACGCAATTAATGTGATTTAGTCTCACTATTAG 1200
Db 1141 CCCGACTGGAAGAGCGGCGAGTGAAGCAACGCAATTAATGTGATTTAGTCTCACTATTAG 1200
QY 1201 GCACAAATCTCATGTTTGAAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGAAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGACAGCATCGGAAGCTGTGTTATGCTGTGAGGCTGTAAATCATTGCATATTCG 1320
Db 1261 TCAGGACAGCATCGGAAGCTGTGTTATGCTGTGAGGCTGTAAATCATTGCATATTCG 1320
```

QY 1321 TGTCCCTCAAGCGGCACTCCCGCTTCGGATATATGTTTTTGGCCGACATCATATACGGTT 1380
DB 1321 TGTCCCTCAAGCGGCACTCCCGCTTCGGATATATGTTTTTGGCCGACATCATATACGGTT 1380
QY 1381 CTGGCAATATTTCTGAATAGCTGTTGACATTAATCATCGGCTCGTATATGTTGTA 1440
DB 1381 CTGGCAATATTTCTGAATAGCTGTTGACATTAATCATCGGCTCGTATATGTTGTA 1440
QY 1441 ATTGTAGCGGATACAAATTTTACACAGGAAACAGCCAGTCCGTTAGTGTTTACGA 1500
DB 1441 ATTGTAGCGGATACAAATTTTACACAGGAAACAGCCAGTCCGTTAGTGTTTACGA 1500
QY 1501 GCATTTACCAACAGGACCATAGATTTATGAAAACTGAAGAGGTAATACTGGTATCTGG 1560
DB 1501 GCATTTACCAACAGGACCATAGATTTATGAAAACTGAAGAGGTAATACTGGTATCTGG 1560
QY 1561 ATTAACGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTAACGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAACTGGAGAGAAATTCGCCAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAACTGGAGAGAAATTCGCCAGGTT 1680
QY 1681 GCGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCACACGACCGTTTGGTGGCTAC 1740
DB 1681 GCGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCACACGACCGTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTGGCTGAATCAACCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
DB 1741 GCTCAATCTGGCCTGTGGCTGAATCAACCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
QY 1801 CGTTTACCTGGATGCGGTAGCTTTACAAACGCAAGCTGATGCTTACCCGATCCTGTT 1860
DB 1801 CGTTTACCTGGATGCGGTAGCTTTACAAACGCAAGCTGATGCTTACCCGATCCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATACAAAGATCTGCTGCCGAACCCGCCAAAACTGGGAA 1920
DB 1861 GAAGCGTTATCGCTGATTTATACAAAGATCTGCTGCCGAACCCGCCAAAACTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGGAAAGGTAGAGCGCGCTGATGTTCAAC 1980
DB 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGGAAAGGTAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGTTATCGCTTCAAG 2040
DB 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGTTATCGCTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGAAGTGGCGTGGATAACCGCTGGCGGAAAGCG 2100
DB 2041 TATGAAAAACGCAAGTACGACATTAAGAAGTGGCGTGGATAACCGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAACCAACACATGATGACGACACCGATTAC 2160
DB 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAACCAACACATGATGACGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTAATAAGGCGAAACAGCGATCACCATCAAGCGCCGCTGG 2220
DB 2161 TCCATCGCAGAGCTGCTTTAATAAGGCGAAACAGCGATCACCATCAAGCGCCGCTGG 2220
QY 2221 GCATGGTCCAAATCGACACCAAGAGTGAATTTATGGTGAACCGTACTGCGGACCTTC 2280
DB 2221 GCATGGTCCAAATCGACACCAAGAGTGAATTTATGGTGAACCGTACTGCGGACCTTC 2280
QY 2281 AAGGTCACCAATCCAAACCGTTGTTGGCTGCTGAGCGCAGGATTAATGAGCGCCAGT 2340
DB 2281 AAGGTCACCAATCCAAACCGTTGTTGGCTGCTGAGCGCAGGATTAATGAGCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGCGAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGCTG 2400
DB 2341 CCGAACAAAGAGCTGCGAAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGCTG 2400

QY 2401 GAAGCGTTAATAAGACAAACCCGCTGGTCCGCTAGCGCTGAAGTCTTACGAGGAAGAG 2460
DB 2401 GAAGCGTTAATAAGACAAACCCGCTGGTCCGCTAGCGCTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCACGTATTGGCCGACCATGGAACACGCCAGAAAGGTGAATCATG 2520
DB 2461 TTGGCGAAAGATCCACGTATTGGCCGACCATGGAACACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCGCGAGATGTCGCTTTCTGGTATGCGGCTGATGCGGTGATCAACGCC 2580
DB 2521 CCGAACATCCGCGAGATGTCGCTTTCTGGTATGCGGCTGATGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCCGACACTAAATTCGAGCTCG 2640
DB 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCCGACACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAAATAACAAATAACAACTCGGATCGAGGGAAGGATTTCAAAATTC 2700
DB 2641 AACAAACAAACAAATAACAAATAACAACTCGGATCGAGGGAAGGATTTCAAAATTC 2700

RESULT 11
AR207294
LOCUS AR207294 7475 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 1 from patent US 6372883.
ACCESSION AR207294
VERSION AR207294.1 GI:21506162
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7475)
AUTHORS Attwood,M.Richard., Hurst,D.Nigel., Jones,P.Stephen.,
Kay,P.Brittain., Raynham,T.Michael. and Wilson,F.Xavier.
TITLE Antiviral medicaments
JOURNAL Patent: US 6372883-A 1 16-APR-2002;
FEATURES Location/Qualifiers
source 1..7475
/organism="unknown"
/mol_type="unassigned DNA"

Query Match 81.8%; Score 2700; DB 6; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGTATGCGCATGATAGCGCCCGGAAGAGA 60
DB 1 CCGACACCATCGAATGGTGCAAAACCTTTCCGGTATGCGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCATTTCCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGACAGTATGCCG 120
DB 61 GTCATTTCCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTCGACAGTATGCCG 120
QY 121 GTGCTCTTTATCAGACCGGTTTCCCGCGTGGTGAACACGAGCCAGCCACGCTTTCTGCGAAA 180
DB 121 GTGCTCTTTATCAGACCGGTTTCCCGCGTGGTGAACACGAGCCAGCCACGCTTTCTGCGAAA 180
QY 181 CGCGGAAAAGTGAAGCGGATGCGGAGCTGAATTAATTAATTCCTCAACCCGCTGGCAC 240
DB 181 CGCGGAAAAGTGAAGCGGATGCGGAGCTGAATTAATTAATTCCTCAACCCGCTGGCAC 240
QY 241 AACAACTCGCGGCAACAGCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300
DB 241 AACAACTCGCGGCAACAGCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300
QY 301 ACGCGCGTTCGCAATTCCTCGCGGATTAATTAATTCGCGCGATCAACTGGGTCGACGG 360
DB 301 ACGCGCGTTCGCAATTCCTCGCGGATTAATTAATTCGCGCGATCAACTGGGTCGACGG 360
QY 361 TGGTGGTTCGATGCTAGAACGAGCGGCTCGAAGCTGTAAAGCGGGGTGCAATC 420
DB 361 TGGTGGTTCGATGCTAGAACGAGCGGCTCGAAGCTGTAAAGCGGGGTGCAATC 420

QY 421 TTCTCGCCAAACGGCTCAGTGGCTGATCAITAACTATCCGCTGATGACCAAGATGCCA 480
DB 421 TTCTCGCCAAACGGCTCAGTGGCTGATCAITAACTATCCGCTGATGACCAAGATGCCA 480
QY 481 TTGCTGTGGAAAGCTCCCTGCCATAATGTTCCGGCGTTATTTCTTGATGCTCTGACGAGA 540
DB 481 TTGCTGTGGAAAGCTCCCTGCCATAATGTTCCGGCGTTATTTCTTGATGCTCTGACGAGA 540
QY 541 CACCCATCAACAGATATTATTTCTCCCATGAAGACGGTACGCACTGGCGCTGGAGCATC 600
DB 541 CACCCATCAACAGATATTATTTCTCCCATGAAGACGGTACGCACTGGCGCTGGAGCATC 600
QY 601 TGGTGGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCCAITTAAGTTTCTGTCTCGG 660
DB 601 TGGTGGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCCAITTAAGTTTCTGTCTCGG 660
QY 661 CGCGTCTCGCTCTGGCTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTCGCTCTGGCTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAAACGGGAAGGACCTGGAGTGCATGTCGGGTTTTCAACAAACCAATGCAAAATGCTGA 780
DB 721 CGGAAACGGGAAGGACCTGGAGTGCATGTCGGGTTTTCAACAAACCAATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGCCAACGATCAGATGGCGCTGGCGGCA 840
DB 781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGCCAACGATCAGATGGCGCTGGCGGCA 840
QY 841 TGGCGGCATTAACGAGTCCGGCTCGCGTGGTGGCGGATATCTCGTAGTGGGATACG 900
DB 841 TGGCGGCATTAACGAGTCCGGCTCGCGTGGTGGCGGATATCTCGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACGCTCATGTTATATCCCGCGGTTAAACACCATCAAAACGAGATTTTC 960
DB 901 ACGATACCGAAGACGCTCATGTTATATCCCGCGGTTAAACACCATCAAAACGAGATTTTC 960
QY 961 GCCTGTGGGGCAAAACAGCGTGAACCGCTTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
DB 961 GCCTGTGGGGCAAAACAGCGTGAACCGCTTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
QY 1021 AGGCAATACGCTGTGGCCCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGGCCAATA 1080
DB 1021 AGGCAATACGCTGTGGCCCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGGCCAATA 1080
QY 1081 CGCAAAACGGCTCTCCCGCGCGTTGGCCGATTCAITTAATGCAGCTGGCAGACAGGTTT 1140
DB 1081 CGCAAAACGGCTCTCCCGCGCGTTGGCCGATTCAITTAATGCAGCTGGCAGACAGGTTT 1140
QY 1141 CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
DB 1141 CCCGACTGGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGATGCGAGTGCACCAATGCTTTCTGGCG 1260
DB 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGATGCGAGTGCACCAATGCTTTCTGGCG 1260
QY 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTGTAATCACTGCATAAATTCG 1320
DB 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGCAGGTGTAATCACTGCATAAATTCG 1320
QY 1321 TGTGCTCAAGCGCATCTCCGTTCTGGATAATGTTTTTGGCGCGACATCATAACCGGTT 1380
DB 1321 TGTGCTCAAGCGCATCTCCGTTCTGGATAATGTTTTTGGCGCGACATCATAACCGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGGA 1440
DB 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTGGA 1440
QY 1441 ATTGTGAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTATAGGTGTTTTACGA 1500
DB 1441 ATTGTGAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTATAGGTGTTTTACGA 1500

QY 1501 GCACCTTCAACAAACGACCATAGATTATGAAAACTGAAAGGTAAACTGGTAACTCTGG 1560
DB 1501 GCACCTTCAACAAACGACCATAGATTATGAAAACTGAAAGGTAAACTGGTAACTCTGG 1560
QY 1561 ATTTAAACGGCGATTAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTTAAACGGCGATTAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTTCCCAACAGGTT 1680
DB 1621 ACCGGAATTAAGGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTTCCCAACAGGTT 1680
QY 1681 GCGGCAACTGCGCATGCGCTGACATTTATCTTCTGGGCAACGACCGCTTTGGTGGCTAC 1740
DB 1681 GCGGCAACTGCGCATGCGCTGACATTTATCTTCTGGGCAACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCCCGGACAAAGGTTTCCAGGACAAAGCTGTAT 1800
DB 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCCCGGACAAAGGTTTCCAGGACAAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACGCGCAAGCTGATTGCTTACCGATCGCTGTT 1860
DB 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACGCGCAAGCTGATTGCTTACCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCCGCCAAAACCTCGGAA 1920
DB 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCCGAACCCGCCAAAACCTCGGAA 1920
QY 1921 GAGATCCCGGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGACGCGCTGATTTCAAC 1980
DB 1921 GAGATCCCGGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGACGCGCTGATTTCAAC 1980
QY 1981 CTGCAAGAAACCGTACTTTCACCTGCGCTGATTGCTGCTCACGGGGGTATGGTTCAAG 2040
DB 1981 CTGCAAGAAACCGTACTTTCACCTGCGCTGATTGCTGCTGACGGGGGTATGGTTCAAG 2040
QY 2041 TATGAAACCGGCAAGTACGACATTAAAGACGCTGGCGTGGATAAACGCTGGCGGAAACG 2100
DB 2041 TATGAAACCGGCAAGTACGACATTAAAGACGCTGGCGTGGATAAACGCTGGCGGAAACG 2100
QY 2101 GGTCTGACCTTCTGTTGACCTGATTAAACAAACACATGATGAGACACCGGATTAC 2160
DB 2101 GGTCTGACCTTCTGTTGACCTGATTAAACAAACACATGATGAGACACCGGATTAC 2160
QY 2161 TCCATCGCAGAAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCGCGCTGG 2220
DB 2161 TCCATCGCAGAAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCGCGCTGG 2220
QY 2221 GCATGFTCCAACTCGACACCGCAAGTGAATTTATGGTGAACGGTACTTCCGACCTTC 2280
DB 2221 GCATGFTCCAACTCGACACCGCAAGTGAATTTATGGTGAACGGTACTTCCGACCTTC 2280
QY 2281 AAGGTCACCAATCCAAACCGTTTCTGTTGGCTGCTGAGCGCAGGTATTAAACGCGCCAGT 2340
DB 2281 AAGGTCACCAATCCAAACCGTTTCTGTTGGCTGCTGAGCGCAGGTATTAAACGCGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGCGCAAAAGGTTTCTCGAAAACTATCTGCTGATGATGAAGTCTG 2400
DB 2341 CCGAAACAAAGAGCTGCGCAAAAGGTTTCTCGAAAACTATCTGCTGATGATGAAGTCTG 2400
QY 2401 GAAGCGGTTTAAATAAGAGCAAAACCGCTGGGTGCGGTAGCGTGAAGTCTTACGAGGAAGAG 2460
DB 2401 GAAGCGGTTTAAATAAGAGCAAAACCGCTGGGTGCGGTAGCGTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGGAAAGATCCAGTATTGGCGGCACCATGGAAAAACCGCCAGAAAGGTGAATCATG 2520
DB 2461 TTGGGAAAGATCCAGTATTGGCGGCACCATGGAAAAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAAACATCCCGCAGATGTCGGCTTTCTGGTATGCGGTGGGTACTGCGGTGATCAAGGCC 2580
DB 2521 CCGAAACATCCCGCAGATGTCGGCTTTCTGGTATGCGGTGGGTACTGCGGTGATCAAGGCC 2580
QY 2581 GCCAGCGGTGCTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640

Db 2581 GCACGGTCTGTCAGACTGTCGATGAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640

Qy 2641 AACAAACAACAATAACAATAACAACACCTCGGGATCGAGGAGGATTTTCAAGATTC 2700

Db 2641 AACAAACAACAATAACAATAACAACACCTCGGGATCGAGGAGGATTTTCAAGATTC 2700

RESULT 12

AX377532

LOCUS AX377532 9191 bp DNA linear PAT 18-MAR-2002

DEFINITION Sequence 9 from Patent WO0212553.

ACCESSION AX377532

VERSION AX377532.1 GI:19573718

KEYWORDS .

SOURCE synthetic construct

ORGANISM synthetic construct

other sequences; artificial sequences.

REFERENCE 1

AUTHORS Kappel,A., Polakowski,T., Pignot,M., Windhab,N., Behrensdoerf,H. and Muth,J.

TITLE Method for detecting mutations in nucleotide sequences

JOURNAL Patent: WO 0212553-A 9 14-FEB-2002;

Nanogen Recognomics GmbH (DE)

FEATURES

source

1. .9191

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

/note="pVALC2x-mutS Plasmid"

ORIGIN

Query Match 81.8%; Score 2698; DB 6; Length 9191;

Best Local Similarity 99.8%; Pred. No. 0;

Mismatches 2701; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 CCGACACCATCAATGGTGCACAAACCTTTCGGGGTATGGCATGATAGCCCGGAAGAGA 60

Db 1 CCGACACCATCAATGGTGCACAAACCTTTCGGGGTATGGCATGATAGCCCGGAAGAGA 60

Qy 61 GTCAATTCAAGGTGGTGAATGTGAACACAGTAACGTTATACGATGTGCGAGAGTATGCCG 120

Db 61 GTCAATTCAAGGTGGTGAATGTGAACACAGTAACGTTATACGATGTGCGAGAGTATGCCG 120

Qy 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCCAGGCCACGCTTTCCTGCGAAAA 180

Db 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCCAGGCCACGCTTTCCTGCGAAAA 180

Qy 181 CCGGGAAAAAGTGGNAGCGCGATGGCGGAGCTGAATTACATTCGCAACCGGTGGGCAC 240

Db 181 CCGGGAAAAAGTGGNAGCGCGATGGCGGAGCTGAATTACATTCGCAACCGGTGGGCAC 240

Qy 241 AACAACTGGCGGCAACAGTCTGCTGATTTGGGGTTCGCCACCTCCAGTCTGGCCCTGC 300

Db 241 AACAACTGGCGGCAACAGTCTGCTGATTTGGGGTTCGCCACCTCCAGTCTGGCCCTGC 300

Qy 301 ACGCGCGTTCGCAAAATTTGTCGGCGGATTAATCTCGCCCGATCAATCGGTCGCCAGCG 360

Db 301 ACGCGCGTTCGCAAAATTTGTCGGCGGATTAATCTCGCCCGATCAATCGGTCGCCAGCG 360

Qy 361 TGGTGTGTCGATGTGTAGAACGAGCGGCTCGAAGCCCTGTAAAGCGCGGTGCACATC 420

Db 361 TGGTGTGTCGATGTGTAGAACGAGCGGCTCGAAGCCCTGTAAAGCGCGGTGCACATC 420

Qy 421 TTCTCGCGCAACCGCTCAGTGGGTGATCATTAATCTATCCGCTGGATGACAGGATGCCA 480

Db 421 TTCTCGCGCAACCGCTCAGTGGGTGATCATTAATCTATCCGCTGGATGACAGGATGCCA 480

Qy 481 TTGCTGTGGAAGTGTGCTGCATTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCGA 540

Db 481 TTGCTGTGGAAGTGTGCTGCATTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCGA 540

Qy 541 CACCCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGGACCTGGCGGTGGAGCATC 600

Db 541 CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600

Qy 601 TGGTGGCATTTGGGTACCAGCAAAATCGCGCTGTTAGCGGGGCCAATTAAGATTTCTGTCTCGG 660

Db 601 TGGTGGCATTTGGGTACCAGCAAAATCGCGCTGTTAGCGGGGCCAATTAAGATTTCTGTCTCGG 660

Qy 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720

Db 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720

Qy 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCACTGCAAAATGCTGA 780

Db 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGTTTTTCAACAAACCACTGCAAAATGCTGA 780

Qy 781 ATGAGGGCATCGTTTCCCACTCGCATGCTGTTGCGCAACGATCAGATGCGCTGGGCGCAA 840

Db 781 ATGAGGGCATCGTTTCCCACTCGCATGCTGTTGCGCAACGATCAGATGCGCTGGGCGCAA 840

Qy 841 TGGCGGCCATTTACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATAG 900

Db 841 TGGCGGCCATTTACCGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATAG 900

Qy 901 ACGATACCGAAGACAGCTCATGTTTATATCCCGCGTTAAACACCATCAAAACAGGATTTTC 960

Db 901 ACGATACCGAAGACAGCTCATGTTTATATCCCGCGTTAAACACCATCAAAACAGGATTTTC 960

Qy 961 GCCTGCTGGGGCAAAACACGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCGCCAGCGGTGA 1020

Db 961 GCCTGCTGGGGCAAAACACGCGTGGACCGCTTGTCTGCAACTCTCTCAGGCGCCAGCGGTGA 1020

Qy 1021 AGGGCAATCAGCTGTGTCCCGTCTCACTGTGTAAGAAAAACAACCCCTGGCGCCCAATA 1080

Db 1021 AGGGCAATCAGCTGTGTCCCGTCTCACTGTGTAAGAAAAACAACCCCTGGCGCCCAATA 1080

Qy 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCTGGCAGCAGAGTTT 1140

Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGCAGCTGGCAGCAGAGTTT 1140

Qy 1141 CCGCATCTGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200

Db 1141 CCGCATCTGAAAGCGGGCAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200

Qy 1201 GCACAAATTTCTATGTTTGCAGCTTATCATGCACTGCAACGCGTGCACCAATGCTTCTGGCG 1260

Db 1201 GCACAAATTTCTATGTTTGCAGCTTATCATGCACTGCAACGCGTGCACCAATGCTTCTGGCG 1260

Qy 1261 TCAGGACGCGCATCGGAAGCTGTGGTATGGCTGTGAGGTCGTAATCACTGCATTAATTCG 1320

Db 1261 TCAGGACGCGCATCGGAAGCTGTGGTATGGCTGTGAGGTCGTAATCACTGCATTAATTCG 1320

Qy 1321 TGTCTCTCAAGCGGCACCTCCCGTCTCGGATTAATGTTTTTGGCGCGACATCATAAACGGTT 1380

Db 1321 TGTCTCTCAAGCGGCACCTCCCGTCTCGGATTAATGTTTTTGGCGCGACATCATAAACGGTT 1380

Qy 1381 CTGGCAAAATTTCTGAAATAGCTGTTGCAATTAATCATCGGCTCGTATTAATGTGTGGA 1440

Db 1381 CTGGCAAAATTTCTGAAATAGCTGTTGCAATTAATCATCGGCTCGTATTAATGTGTGGA 1440

Qy 1441 ATTGTAGCGGATTAACAAATTTTCAACAGGAACAGCCAGTCCGTTTAGGTGTTTTCAGA 1500

Db 1441 ATTGTAGCGGATTAACAAATTTTCAACAGGAACAGCCAGTCCGTTTAGGTGTTTTCAGA 1500

Qy 1501 GCACTTTCAACCAACAGGACCATAGCATATGAATAATCGAAGAGGTAACTGGTATCTGG 1560

Db 1501 GCACTTTCAACCAACAGGACCATAGCATATGAATAATCGAAGAGGTAACTGGTATCTGG 1560

Qy 1561 ATTAACGGCGATAAAGGCTATAACCGGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620

Db 1561 ATTAACGGCGATAAAGGCTATAACCGGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620

Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAACTGGAAGAGAAATTCGCCACAGTTT 1680

Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAAATCCCAACAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGGCTGTGTGGCTGAAATCAACCCCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGGCTGTGTGGCTGAAATCAACCCCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCGGTAGCTTACACGCAAGCTGATGCTTACCGATCCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCGGTAGCTTACACGCAAGCTGATGCTTACCGATCCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTTATACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGCAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGCAA 1920
QY 1921 GAGATCCGCGCTGGATAAGAACTGAAAGCGAAAGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCGCGCTGGATAAGAACTGAAAGCGAAAGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCTGACTTACCTGGCGCTGATGCTGCTGACGGGGTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCTGACTTACCTGGCGCTGATGCTGCTGACGGGGTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACATGAATGCAAGACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACATGAATGCAAGACCGGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAATAAGCGCAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAATAAGCGCAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCAAATCGACACCAAGCAAGTGAATATGCTGTAACCGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACCAAGCAAGTGAATATGCTGTAACCGTACTGCCGACCTTC 2280
QY 2281 AAGGTCACACCATCAACCGCTTGGTGGCTGCTGACGCGCAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGTCACACCATCAACCGCTTGGTGGCTGCTGACGCGCAGGTATTAACGCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACATCTGCTGACTGATGAAGGCTCG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACATCTGCTGACTGATGAAGGCTCG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTCCGCTAGCGCTGAACTTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTCCGCTAGCGCTGAACTTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCACGCTATGCGCCACCATGGAACCGCCGCAAGGTAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGCTATGCGCCACCATGGAACCGCCGCAAGGTAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCGTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACCAACACCAATAAACAATAACAAACCTCGGGATCGAGGGAAGATTTTCAGATTC 2700
Db 2641 AACCAACACCAATAAACAATAACAAACCTCGGGATCGAGGGAAGATTTTCAGATTC 2700
QY 2701 GGATCC 2706
Db 2701 GGATCC 2706

RESULT 13
AF097412
LOCUS
DEFINITION
AF097412
ACCESSION
AF097412.1 GI:3983120
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 6724
/organism="Expression vector MBP-parallel1"
/mol_type="other DNA"
/db_xref="taxon:84479"
204. .1163
/gene="lacIq"
204. .1163
/gene="lacIq"
/codon_start=1
/transl_table=11
/product="lac repressor"
/protein_id="AAC83812.1"
/db_xref="GI:3983121"
translation="MAELNYIPNRVAQLAGKQSLIGVATSLALHAPSOIVAAIKS
RAQLGASVVSVMSRSGVEACKAAVHNLQAQVGLIINYPDDQDAIAVEACATNV
PAFLDVSQTPINSIIFSHEDQTRGLVHELVAGHQIALLAGPLSSARLRLAGW
HKYLRNQIQAEREGDWSAMSGFOQTQMLNEGIVPTAMLVANDQMALGAMRAITE
SGLRVADISVVGVDDESDSSCYIPPLTTIKODFRLLGQTSVDRLLQLSQGQAVKGNQ
LLPVSUVKRTTLAPNTQTASPRALADSLMQLARQVSRLESQ"
1453. .1481
/note="tac; inducible promoter"
1528. .2994
/gene="male"
1528. .2994
/gene="male"
/codon_start=1
/transl_table=11
/product="maltose binding protein"
/protein_id="AAC83813.1"
/db_xref="GI:3983122"
translation="WKIEBGLVINGDKYGLAEVKGKPEKDTGIKVTVHEHPDKL
EKKFPOVATGDPDIIFWAHDRCGYAGSGLLAEITPKAFQDKLYPFTWDAVRNG
KLIAVDIAVEALSLIYNKDLLPNPKTWEEIPALDKELKAKGSAFLNQLQBPFTWP
LIAADGGYAFKTEGNGKIDKGVONAGAKAGTLFVLDLTKNKHMDYDLSIAEAF
NKGEMTINGFWANSIDTSKVNYGTVLTFKQKQSPKFPVGLSAGTNAASPNKEL
AKKEFLNYLLTDEGLEAVNKKDKPLGAKVLSYEELAKDPRIAAATMENAKGGEIMENI
PQMSAFYAVRTAVINAASGROTVDBALKDAOTNSSNNNNNNNNNNLIGDITENLYF
QGAMPDFPKLRRRLQVRLPLNLRPAVSRHAPVSLALAVLQRRDWPENPGVTQLNRL
AAHPPASWNSSEARTDRPSQQLSLNGEWQLGCEFG"
2633. .2688
/gene="male"
/note="protein spacer region"
2689. .2709
/gene="male"
/note="Tobacco etch virus protease cleavage site"
2711. .2810
/gene="male"
misc_feature
misc_feature
misc_feature

gene	/note="multiple cloning site"				
promoter	3426..4356			Db	721
	/gene="bla"				
	3426..3454			Qy	781
	/gene="bla"				
CDS	3496..4356			Db	781
	/gene="bla"				
	/codon_start=1			Qy	841
	/transl_table=11				
	/product="beta-lactamase"			Db	841
	/protein_id="AAC83814.1"				
	/db_xref="GI:3983123"			Qy	901
	/translation="MSIOHPRVALIPFAARCLPVFAHPETLVKVKDAEDQLGARVGY IELDNLAKILLESFRRPERSPMWTFKVLICGAVLSRYDAGBQELRRIRHYSDNLVE YSPVTEKHLDGMIVRELCSAAITWSDINLANLLTTTGGPKELTAFILHNMGDHDVRL DRWPELNEAIIPNDRDTPVAMATTLRLKLTIGELLTLASRQQLIDWMEADKVGAPL LRSLAPAGWFIADKSGAGERSGRIIAALGPDGKPSRIWVITVITGSOATMDERNQIA BIGASLIKHW"			Db	901
ORIGIN					
Query Match	81.1%;	Score 2676.2;	DB 12; Length 6724;		
Best Local Similarity	99.9%;	Pred. No. 0;			
Matches 2678;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;	
Qy	1	CGGACACCATCGAATGGTGC	AAAACCTTTCGGGTATGGCATGATAGCCGCGCGGAAGAGA	60	
Db	1	CGGACACCATCGAATGGTGC	AAAACCTTTCGGGTATGGCATGATAGCCGCGCGGAAGAGA	60	
Qy	61	GTCAATTCAGGCGTGGTGAATGTGA	AAACCACTTATACCGATGTCGCAGAGTATGCGG	120	
Db	61	GTCAATTCAGGCGTGGTGAATGTGA	AAACCACTTATACCGATGTCGCAGAGTATGCGG	120	
Qy	121	GTGTCCTTTATCAGACCGTTTCCG	CGGTGTAACCAAGCCACGCGTTTCTGCGAAAA	180	
Db	121	GTGTCCTTTATCAGACCGTTTCCG	CGGTGTAACCAAGCCACGCGTTTCTGCGAAAA	180	
Qy	181	CGCGGAAAGTGGAGCGGCGATGG	CGGAGCTGAATTTACATTTCCCAACCCGCGTGGCAC	240	
Db	181	CGCGGAAAGTGGAGCGGCGATGG	CGGAGCTGAATTTACATTTCCCAACCCGCGTGGCAC	240	
Qy	241	AACAACTGCGCGGCAAAACAGTGT	GCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300	
Db	241	AACAACTGCGCGGCAAAACAGTGT	GCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300	
Qy	301	ACGCGCCGTGCGCAAAATTTGTCG	CGGCGATTAATAATCTCGCGCGGATCAACTGGGTGCCAGCG	360	
Db	301	ACGCGCCGTGCGCAAAATTTGTCG	CGGCGATTAATAATCTCGCGCGGATCAACTGGGTGCCAGCG	360	
Qy	361	TGGTGGTGCATGTGTAGAACCAAG	CGCGCTCGAAGCCCTGTAAAGCGCGGTGCACAATC	420	
Db	361	TGGTGGTGCATGTGTAGAACCAAG	CGCGCTCGAAGCCCTGTAAAGCGCGGTGCACAATC	420	
Qy	421	TTCTCGCGCAACCGCTCAGTGGGCT	GATCATTTAACTATCCGCTGGATGACCAAGGATGCCA	480	
Db	421	TTCTCGCGCAACCGCTCAGTGGGCT	GATCATTTAACTATCCGCTGGATGACCAAGGATGCCA	480	
Qy	481	TTGCTGTGGAAGCTGCCTGC	ACTAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCCAGA	540	
Db	481	TTGCTGTGGAAGCTGCCTGC	ACTAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCCAGA	540	
Qy	541	CACCCATCAACAGTATTATTTCT	CCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600	
Db	541	CACCCATCAACAGTATTATTTCT	CCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600	
Qy	601	TGGTGCATTTGGGTACACAGAAAT	CGCGCTGTAGCGGCGCCATTAAGTCTGTCTCGG	660	
Db	601	TGGTGCATTTGGGTACACAGAAAT	CGCGCTGTAGCGGCGCCATTAAGTCTGTCTCGG	660	
Qy	661	CGCGTCTGCGTCTGGCTGGCATA	AAATATCTCCTCAGTCAATCAAAATTCAGCCGATAG	720	
Db	661	CGCGTCTGCGTCTGGCTGGCATA	AAATATCTCCTCAGTCAATCAAAATTCAGCCGATAG	720	
Qy	721	CGGAACGGGAAGGCGACTGGAGT	GCGATGTCCGGTTTTTCAACAACCAATGCAATGCTGA	780	

Db 1801 CCGTTTACCTGGAGTCCGTA CGTTACAACGCAAGCTGATTCTTACCCGATCGCTGT 1860
Qy 1861 GAAGCGTTATCGTGTATTTATAACAAGATCTGCTGCCGAACCGCCAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGTGTATTTATAACAAGATCTGCTGCCGAACCGCCAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGSTTAAGAGCGCGTGTATTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGSTTAAGAGCGCGTGTATTTCAAC 1980
Qy 1981 CTGCAAGAACCTACTTACCTGGCGCTGATTGCTGTGACGGGGTTATCGCTTCAAG 2040
Db 1981 CTGCAAGAACCTACTTACCTGGCGCTGATTGCTGTGACGGGGTTATCGCTTCAAG 2040
Qy 2041 TATGAAAACGCGCAAGTACGACATTAAGACGTGGCGTGGATAACCGTGGCGGAAACG 2100
Db 2041 TATGAAAACGCGCAAGTACGACATTAAGACGTGGCGTGGATAACCGTGGCGGAAACG 2100
Qy 2101 GGTCTGACCTTCTCGTTGACCTGATTAAACCAACACATGAATGACAGACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAAACCAACACATGAATGACAGACCGGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCCAAATCGACACACGCAAGTGAATTAATGTTGAACGGTACTCGCCACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACACGCAAGTGAATTAATGTTGAACGGTACTCGCCACCTTC 2280
Qy 2281 AAGGTCACACATCAACACCGTTCTGTTGGTGTCTGAGCGCAGGTATTAACGGCCCGAGT 2340
Db 2281 AAGGTCACACATCAACACCGTTCTGTTGGTGTCTGAGCGCAGGTATTAACGGCCCGAGT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAACATCTCTGACTGATGAAGGCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAACATCTCTGACTGATGAAGGCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGTCCGTTAGCGCTGAAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTCCGTTAGCGCTGAAAGTCTTACGAGGAAG 2460
Qy 2461 TTGGCGAAGATCCACGTTATTCGCCACACATGAAACGCCCAAGAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGTTATTCGCCACACATGAAACGCCCAAGAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGCTCCGCTTTCTGTTATGCGGTGCTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGCTCCGCTTTCTGTTATGCGGTGCTACTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACATGCTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACATGCTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACAAACAAACAAATAACAATAACAACCTCGGGATCGA 2681
Db 2641 AACAAACAAACAAATAACAATAACAACCTCGGGATCGA 2681

RESULT 14
AX378208
LOCUS AX378208 6806 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 194 from Patent WO0206306.
ACCESSION AX378208
VERSION AX378208.1 GI:19574061
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
AUTHORS Yan, R., Tomasselli, A. G., Gurney, M. E., Emmons, T. L., Bienkowski, M. J.
and Heinrikson, R. L.
TITLE Substrates and assays for (g)-secretase activity

JOURNAL Patent: WO 0206306-A 194 24-JAN-2002;
PHARMACIA & UPJOHN COMPANY (US)
FEATURES Location/Qualifiers
source 1..6806
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
/note="fusion protein comprising a maltose binding protein
with 125 amino"

ORIGIN
Query Match 79.9%; Score 2636.8; DB 6; Length 6806;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2638; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 CCGACACATCGAATGTGCAAAACCTTTCCGGTATGGCATGATAGCGCCGGAAGAGA 60
Db 1 CCGACACATCGAATGTGCAAAACCTTTCCGGTATGGCATGATAGCGCCGGAAGAGA 60
Qy 61 GTCAATTCAGGTGTGAATGTGAACACAGTAACTTATACGATGTCGAGAGTATGCCG 120
Db 61 GTCAATTCAGGTGTGAATGTGAACACAGTAACTTATACGATGTCGAGAGTATGCCG 120
Qy 121 GTGCTCTTATCAGACCGCTTTCCCGCTGTGAACACAGGCGACCGTTCCTGCGAAA 180
Db 121 GTGCTCTTATCAGACCGCTTTCCCGCTGTGAACACAGGCGACCGTTCCTGCGAAA 180
Qy 181 CGCGGAAAAGTGAAGCGCGATGCGGAGCTGAATTAATTCACATTCGCAACCGCGTGCAC 240
Db 181 CGCGGAAAAGTGAAGCGCGATGCGGAGCTGAATTAATTCACATTCGCAACCGCGTGCAC 240
Qy 241 AACAACTGGCGGCAACACAGTCTGCTGATTCCGCTTGGCCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAACACAGTCTGCTGATTGGCGTTGCCCTCCAGTCTGGCCCTGC 300
Qy 301 ACGCGCGCTCGCAAAATTGTGCGCGCGATTAAATTCCTGCGCGGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGCTCGCAAAATTGTGCGCGCGATTAAATTCCTGCGCGGATCAACTGGGTGCCAGCG 360
Qy 361 TGTGCTGTGCTGATGATGAACGAGCGGCTGCAAGCTGTAAAGCGGCGGTGCAAAATC 420
Db 361 TGTGCTGTGCTGATGATGAACGAGCGGCTGCAAGCTGTAAAGCGGCGGTGCAAAATC 420
Qy 421 TTCTCGCGCAACCGCTCAGTGGCTCATTAATTAATTCCTGCGCGGATCAACTGGGTGCCAG 480
Db 421 TTCTCGCGCAACCGCTCAGTGGCTCATTAATTAATTCCTGCGCGGATCAACTGGGTGCCAG 480
Qy 481 TTGCTGTGGAAGCTGCTGCACTAAATGTTCCGCGCTTATTTCTTGATGTCTCTGACCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAAATGTTCCGCGCTTATTTCTTGATGTCTCTGACCAGA 540
Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAGACGCTAGCGACTGGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAGACGCTAGCGACTGGGCGGTGGAGCATC 600
Qy 601 TGTGCTGCTTGGCTCACAGCAAAATCGCGCTGTTAGCGGGGCCCATTAAGTCTGCTCGG 660
Db 601 TGTGCTGCTTGGCTCACAGCAAAATCGCGCTGTTAGCGGGGCCCATTAAGTCTGCTCGG 660
Qy 661 CGGCTCTGCTGTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGGCTCTGCTGTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACCGGAAGCGCACTGGAGTGCCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACCGGAAGCGCACTGGAGTGCCATGTCGGTTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGCAATCGTTCCTACTCGCATGCTGGTTGCCAACGATCAGATGGCGCTGGCGCA 840
Db 781 ATGAGGCAATCGTTCCTACTCGCATGCTGGTTGCCAACGATCAGATGGCGCTGGCGCA 840
Qy 841 TCGCGCGCAATACCGAGTCCGGCTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db 841 TCGCGCGCAATACCGAGTCCGGCTGGTGGCGATATCTCGGTAGTGGGATACG 900

Db 841 TCGCGCCATTACGAGTCCGGGCTGCGGTTGGTGGGATATCTCGTAGTGGATAGG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCGCCCGTTAAACCAATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCGCCCGTTAAACCAATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
QY 1021 AGGCGAATCAGCTGTTGCCCGCTCACTCGTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGCGAATCAGCTGTTGCCCGCTCACTCGTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGCTGGCGGATTCATTAATGACGTGGCAGCAGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGCTGGCGGATTCATTAATGACGTGGCAGCAGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGACGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGACGCAACGCAATTAATGTGAGTTAGCTCACTCATTAG 1200
QY 1201 GCACAATCTCATGTTTGACAGCTTATCATCACTGCAGGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATCTCATGTTTGACAGCTTATCATCACTGCAGGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGAGCCATCGGAAGCTGTGATGGCTGTGAGGTGCTGAATCACTGATTAATTCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGATGGCTGTGAGGTGCTGAATCACTGATTAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACTCCCGTTCTGATATGTTTGTGGCGGACATCATACAGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCTGATATGTTTGTGGCGGACATCATACAGGTT 1380
QY 1381 CTGGCAATATCTGAAATGAGCTGTGACATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAATATCTGAAATGAGCTGTGACATTAATCATCGGCTCGTATATGTTGGA 1440
QY 1441 ATTGTGAGCGGATAACAATTTACACAGGAACAGCCAGTCCGTTTGTGTTTTCAGGA 1500
Db 1441 ATTGTGAGCGGATAACAATTTACACAGGAACAGCCAGTCCGTTTGTGTTTTCAGGA 1500
QY 1501 GCACCTCACCAACAGGACCATAGATTATGAAGAACTGAAGAGGTAACCTGTTATCTGG 1560
Db 1501 GCACCTCACCAACAGGACCATAGATTATGAAGAACTGAAGAGGTAACCTGTTATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACGCTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGCTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGAATAAATTCGGAAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGAATAAATTCGGAAGAAATTCACAGGTT 1680
QY 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACGACCGCTTGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACGACCGCTTGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGGCTGAATCACTCCCGGCAAGCGTTCCAGACAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAATCACTCCCGGCAAGCGTTCCAGACAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCGCTAGCTTACACGCAAGCTGATTGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGATGCGCTAGCTTACACGCAAGCTGATTGCTTACCCGATCGCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTTTATACAAAGATCTGCTGCCGAACCCGCCAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTTATACAAAGATCTGCTGCCGAACCCGCCAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAGAGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAGAGCGCTGATGTTCAAC 1980

QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGTGACGCGGGTTATATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGTGACGCGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAGAACGCAAGTACGACATTAAGACGCTGGCGGTGGATTAAGCTGGCGCGAAGCG 2100
Db 2041 TATGAAGAACGCAAGTACGACATTAAGACGCTGGCGGTGGATTAAGCTGGCGCGAAGCG 2100
QY 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAGAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAAGAAACAAACACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGAGAAAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACGCGCCGTGG 2220
Db 2161 TCCATCGAGAAAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACGCGCCGTGG 2220
QY 2221 GCATGCTCAACATCGACACGCAAGCAAGTGAATTAATGTGTAAACGCTGTCGCACTTC 2280
Db 2221 GCATGCTCAACATCGACACGCAAGCAAGTGAATTAATGTGTAAACGCTGTCGCACTTC 2280
QY 2281 AAGGCTCAACCATCCAAACCGTTCTGTTGGCGTGTGAGCGAGGTATTAAGCGCCCGCAGT 2340
Db 2281 AAGGCTCAACCATCCAAACCGTTCTGTTGGCGTGTGAGCGAGGTATTAAGCGCCCGCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGTTAATAAGACAAACCGCTGGTCCGTAGCGCTGAGCTTACGAGGAAG 2460
Db 2401 GAAGCGTTAATAAGACAAACCGCTGGTCCGTAGCGCTGAGCTTACGAGGAAG 2460
QY 2461 TTGGCAAGAGTCCACGATTTTGGCGCCACCACTGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCAAGAGTCCACGATTTTGGCGCCACCACTGGAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACAATCCGCGAGATGTCGCTTTTGTGATGCCGTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACAATCCGCGAGATGTCGCTTTTGTGATGCCGTGCTGCTGATCAACGCC 2580
QY 2581 GCCAGCGTCTCAGACTGCTGATGAAGCCCTGGAAGCGCGCAGACTTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTCAGACTGCTGATGAAGCCCTGGAAGCGCGCAGACTTAATTCGAGCTCG 2640

RESULT 15

AF031088 6706 bp DNA circular SYN 07-MAY-2001
LOCUS Shuttle vector pMAL-pIII, complete sequence.
DEFINITION AF031088
ACCESSION AF031088
VERSION AF031088.1 GI:2623821

KEYWORDS Shuttle vector pMAL-pIII

SOURCE Shuttle vector pMAL-pIII

ORGANISM Shuttle vector pMAL-pIII

OTHER SEQUENCES; artificial sequences; vectors.

REFERENCE 1 (bases 1 to 6706)

Zwick, M.B., Bonnycastle, L.L., Noren, K.A., Venturini, S., Leong, E.,
Barbas, C.F. III, Noren, C.J. and Scott, J.K.
The maltose-binding protein as a scaffold for monovalent display of
peptides derived from phage libraries
Anal. Biochem. 264 (1), 87-97 (1998)

JOURNAL MEDLINE 99002881

PUBMED 9784192

REFERENCE 2 (bases 1 to 6706)

Noren, K.A. and Noren, C.J.

Direct Submission

Submitted (22-OCT-1997) New England Biolabs, 32 Tozer Road,

Beverly, MA 01915, USA

Location/Qualifiers

1..6706

source

/organism="Shuttle vector pMAL-pIII"

/mol_type="genomic DNA"

```

/db_xref="taxon:69327"
/note="derived from commercially available vector pMal-p2
(New England Biolabs)"
15..1163
/gene="lacI"
/note="E. coli gene"
15..46
/gene="lacI"
81..1163
/gene="lacI"
/note="transcriptional repressor"
/codon_start=1
/transl_table=11
/product="lac repressor"
/protein_id="AAB86558.1"
/db_xref="GI:2623822"
/translation="MKPVTLYDVARYAGVSYQTVSRVUNQASHVSAKTRKVEAMAB
LNYIPNRVAQLAGQSLIGVATSSLAHAFPSQIVAAIKRADQLGASVVSVSERS
GVEACKAAPHLLAQVRVGLIINYPDLDDQDAIEAEACTNVPALFLVDSDQPTINSII
FSHEDGRLGVEHLVALGQIALLAGLSVSARLRAGWHKYLTRNQIDQPIAEREG
DWSAMSGFQQTQMQLNEGIVPTAMLVANDQMALGAMRAITESGLRVGADI SVGVYDDT
EDSCYIPPLTTIKQDFRLGQTSVDRLLQSQQAVKGNQLLPVSLVKRKTLLPNT
QTASPRALADSLMLQARQVSRLESQ"
1406..1433
/note="PtaC promoter"
1528..2976
/gene="malE-lacZa fusion"
/note="E. coli gene fusion"
1528..2976
/gene="malE-lacZa fusion"
/note="MBP-lacZa; MBP modified to include phage M13
secretion signal sequence"
/codon_start=1
/transl_table=11
/product="maltose binding protein-lacZ alpha peptide
fusion protein precursor"
/protein_id="AAB86559.1"
/db_xref="GI:2623823"
/translation="MKKLLFAIPLVVPFYSHSADMKIEBGLKVIWINDKGYNGLAHV
GKFEQDTGIVTVHEPDKLEEKPPQVAATGDGPDII FWAHDFPGVAGQLLAEITP
DKAFQDKLYPTDVAHYNGKLIAYPIAVEALSIIYNKOLLPNPKTWBEI PALDKEL
KAGKSALMFLNPQPYFTPLIAADGGVAFKGVKVDI KDVGVDNAGAKAGLTFLDV
LKNKHNADTDYSIAEAPNKGETAMTINGPWNASNDITKVNYGVTVLPTFGQVDS
KPFYGVLSAGTNAASPENKELAKFENLYLIDEGLEAVNKDKPLGVALVKSBEELAK
DPRIAATHENAKQIEIPNIPQMSAFWIAVTRVINAASGRQTVDEALDKDAQTNSSN
NNNNNNNLGIEIRIIEFGSSRVLDQASLALAVLQRRDWNENPGVTQLNRLAAHPFP
ASWRNSEARTDRPSQQLRSLSNGEWQLGCFGG"
1528..1581
/gene="malE-lacZa fusion"
/note="M13 secretion signal sequence"
1582..2973
/gene="malE-lacZa fusion"
/product="maltose binding protein-lacZ alpha peptide
fusion protein"
3478..4338
/gene="bla"
/note="R-plasmid ampicillin resistance gene"
3478..4338
/gene="bla"
/note="penicillinase"
/codon_start=1
/transl_table=11
/product="RTEW-1 beta-lactamase"
/protein_id="AAB86560.1"
/db_xref="GI:2623824"
/translation="MSIQHPRVRLIIPFFAAFCPLPVFAHPETLVKVKDAEDQLGARVGY
IELDLNSKILLESFRRPEERPMWMTFKVLLCGAVLSRVDAAGQQLGRRIRHYSDNLAVE
YSPVTEKHLTDGMTVERELCSAAITMSDNTAANLLITIGPKELTAFLHMGDHTVRL
DRWEPELNEAI PNDERDTPVAMATTLRKLLTGELLTLASRQQLIDWMEADKVAQPL
LRGALPAGWFIADKSGNGERGSGRIIAALGFDGKPSRI VVIYITGTSQATMDERNROIA
EIGASLIKHW"
4381..4893
/standard_name="M13 origin (+ -)"

```

```

ORIGIN
Query Match      79.7%; Score 2631; DB 12; Length 6706;
Best Local Similarity 97.7%; Pred. No. 0;
Matches 2704; Conservative 0; Mismatches 5; Indels 60; Gaps 1;

QY 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGCGGTATGGCATGATAGGCCCGGAAAGAGA 60
DB 1 CCGACACCATCGAATGGTGC AAAACCTTTTCGCGGTATGGCATGATAGGCCCGGAAAGAGA 60

QY 61 GTC AATT CAGGGTGGTGAATGTGA AACCAAGTAACTGTTATACGATGTGCGAGAGTATCCG 120
DB 61 GTC AATT CAGGGTGGTGAATGTGA AACCAAGTAACTGTTATACGATGTGCGAGAGTATCCG 120

QY 121 GTG TCTCTTATCAGACCGTTTCCGCGTGTGAACCAAGGCCAGCCAGTTTCTTGC GAAAA 180
DB 121 GTG TCTCTTATCAGACCGTTTCCGCGTGTGAACCAAGGCCAGCCAGTTTCTTGC GAAAA 180

QY 181 CGCGGGAAAAAGTGGAAAGCGCGATGCGGAGCTGGAATTACATTCCCAACCGCGTGGCAC 240
DB 181 CGCGGGAAAAAGTGGAAAGCGCGATGCGGAGCTGGAATTACATTCCCAACCGCGTGGCAC 240

QY 241 AAC AACTGCGCGGCAACACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
DB 241 AAC AACTGCGCGGCAACACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300

QY 301 ACGGCGCGTCGCAAAATTTGTCGCGCGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
DB 301 ACGGCGCGTCGCAAAATTTGTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 360

QY 361 TGTGTGTTGTCGATGTTAGAACGAAGCGGCTGCAAGCCCTGTAAGCGCGGTGCACAATC 420
DB 361 TGTGTGTTGTCGATGTTAGAACGAAGCGGCTGCAAGCCCTGTAAGCGCGGTGCACAATC 420

QY 421 TTTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACCAAGTGCCA 480
DB 421 TTTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACCAAGTGCCA 480

QY 481 TTGCTGTGGAAGCTGCTGCACCTAAATTTTCGCGCGTATTTCTTGATGTCTCTGACGAGA 540
DB 481 TTGCTGTGGAAGCTGCTGCACCTAAATTTTCGCGCGTATTTCTTGATGTCTCTGACGAGA 540

QY 541 CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGGACCTGGCGGTGGAGCATC 600
DB 541 CACCCATCAACAGTATATTTTCTCCCATGAAGACGGTACGGACCTGGCGGTGGAGCATC 600

QY 601 TGTGTGCATTTGGTTCACGAGCAATCCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
DB 601 TGTGTGCATTTGGTTCACGAGCAATCCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660

QY 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720

QY 721 CGGAACGGGAAGCGCATCGAGTGC CATGTCGGTTTTCAACAAACCATGCATGCTGA 780
DB 721 CGGAACGGGAAGCGCATCGAGTGC CATGTCGGTTTTCAACAAACCATGCATGCTGA 780

QY 781 ATGAGGGCATCGTTCCTCACTCGCATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840
DB 781 ATGAGGGCATCGTTCCTCACTCGCATGCTGGTTGCCAACGATCAGATGGCGCTGGGCGCAA 840

QY 841 TGGCGGCCATTACCGAGTCCGCGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATACG 900
DB 841 TGGCGGCCATTACCGAGTCCGCGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATACG 900

QY 901 ACGATACCGAAGACAGCTCATGTTTATATCCCGCGTTTAAACCACTCAAAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTTATATCCCGCGTTTAAACCACTCAAAACAGGATTTTC 960

QY 961 GCCTGCTGGGGCAAAACCAAGCGTGGACCGCTTCTGCTCAACTCTCTCAGGGCCAGGGGTGA 1020
DB 961 GCCTGCTGGGGCAAAACCAAGCGTGGACCGCTTCTGCTCAACTCTCTCAGGGCCAGGGGTGA 1020

```

QY	1021	AGGCAATCAGCTGTGCCGCTCTCACTGGTGAAAGAAAAACCAACCTGGCGCCCAATA	1080
Db	1021	AGGCAATCAGCTGTGCCGCTCTCACTGGTGAAAGAAAAACCAACCTGGCGCCCAATA	1080
QY	1081	CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGACGTGGCAGCAGAGTTT	1140
Db	1081	CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGACGTGGCAGCAGAGTTT	1140
QY	1141	CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG	1200
Db	1141	CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG	1200
QY	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCGAGGTGCACCAATGCTTCTGGCG	1260
Db	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCGAGGTGCACCAATGCTTCTGGCG	1260
QY	1261	TCAGGACGCATCGGAAGCTGTGGTATGCTGTGCGAGGTGTAATCACTGCATAAATTCG	1320
Db	1261	TCAGGACGCATCGGAAGCTGTGGTATGCTGTGCGAGGTGTAATCACTGCATAAATTCG	1320
QY	1321	TGTCGCTCAAGCGCACTCCCGTCTCGGATAATGTTTTTGGCGCGACATCAATACGGTT	1380
Db	1321	TGTCGCTCAAGCGCACTCCCGTCTCGGATAATGTTTTTGGCGCGACATCAATACGGTT	1380
QY	1381	CTGGCAATATCTGAATAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGTGA	1440
Db	1381	CTGGCAATATCTGAATAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGTGA	1440
QY	1441	ATTGTGACGGATACAAATTTACACAGCAAGCAGCTCGTGTAGTGTTTTACGA	1500
Db	1441	ATTGTGACGGATACAAATTTACACAGCAAGCAGCTCGTGTAGTGTTTTACGA	1500
QY	1501	GCCTTCCACCAACAGGACCATAG-----	1524
Db	1501	GCCTTCCACCAACAGGACCATAGCATATGAAAAATTTATTTCGCAATTCCTTTAGTG	1560
QY	1525	-----ATTATGAAAACTGAAGAGGTAAACTGTAATCTGG	1560
Db	1561	GTACCTTTCTATCTCACTCGGCCGATATGAAAAATCGAAGAGGTAAACTGTAATCTGG	1620
QY	1561	ATTTACGCGATAAAGGCTATAACGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
Db	1621	ATTTACGCGATAAAGGCTATAACGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1680
QY	1621	ACCGAATTTAAAGTCAACCGTTGAGCATCCGATAAACTGGAAGAGAAATTCGCCAGGTT	1680
Db	1681	ACCGAATTTAAAGTCAACCGTTGAGCATCCGATAAACTGGAAGAGAAATTCGCCAGGTT	1740
QY	1681	GGGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
Db	1741	GGGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1800
QY	1741	GCTCAATCTGGCCTGTGGCTGAATCAACCGGCAAGCGTTCCAGGCAAGCTGTAT	1800
Db	1801	GCTCAATCTGGCCTGTGGCTGAATCAACCGGCAAGCGTTCCAGGCAAGCTGTAT	1860
QY	1801	CGGTTTACCTGGGATGCGGCTAGCTTTACACGCGCAAGCTGATTGCTTACCGGATCGCTGTT	1860
Db	1861	CGGTTTACCTGGGATGCGGCTAGCTTTACACGCGCAAGCTGATTGCTTACCGGATCGCTGTT	1920
QY	1861	GAAGCGTTATCGCTGATTTTAAAGAGATCTGCTGCGGAAACCGGCCAAAAACCTGGGAA	1920
Db	1921	GAAGCGTTATCGCTGATTTTAAAGAGATCTGCTGCGGAAACCGGCCAAAAACCTGGGAA	1980
QY	1921	GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAGGTGAAGCGCGCTGATGTTCAAC	1980
Db	1981	GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAGGTGAAGCGCGCTGATGTTCAAC	2040
QY	1981	CTGCAAGAACCGTACTTCACTGGCGGCTGATTGCTGTGACGGGGTTTATCGGTTCAAG	2040
Db	2041	CTGCAAGAACCGTACTTCACTGGCGGCTGATTGCTGTGACGGGGTTTATCGGTTCAAG	2100

QY	2041	TATGAACCGCAAGTACGACATTTAAAGACGTGGCGTGGATAACGCTGGCGCGAAGCG	2100
Db	2101	TATGAACCGCAAGTACGACATTTAAAGACGTGGCGTGGATAACGCTGGCGCGAAGCG	2160
QY	2101	GGTCTGACCTTCTCTGGTTGACCTGATTAATAAAACAAACATGAAATGCAGACACCGATTAC	2160
Db	2161	GGTCTGACCTTCTCTGGTTGACCTGATTAATAAAACAAACATGAAATGCAGACACCGATTAC	2220
QY	2161	TCCATCGCAGAAAGTGCCTTTAATAAAGCGAAACAGCGATGACCATCAACGGCCCGTGG	2220
Db	2221	TCCATCGCAGAAAGTGCCTTTAATAAAGCGAAACAGCGATGACCATCAACGGCCCGTGG	2280
QY	2221	GCATGCTCCACATCGACACGACGAAAGTGAATTAATGGTGTAAACGGTACTCGGACCTTC	2280
Db	2281	GCATGCTCCACATCGACACGACGAAAGTGAATTAATGGTGTAAACGGTACTCGGACCTTC	2340
QY	2281	AAGGTCAACCATCCAAACCGTTTGGCGTGTGAGCGAGGTATTAAGCGCCCGCAGT	2340
Db	2341	AAGGTCAACCATCCAAACCGTTTGGCGTGTGAGCGAGGTATTAAGCGCCCGCAGT	2400
QY	2341	CCGAACAAAGAGTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCTG	2400
Db	2401	CCGAACAAAGAGTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCTG	2460
QY	2401	GAAGGGTTAATAAGACAAACCGCTGGTCCGTAGCGCTGAAAGTCTTACGAGGAAGAG	2460
Db	2461	GAAGGGTTAATAAGACAAACCGCTGGTCCGTAGCGCTGAAAGTCTTACGAGGAAGAG	2520
QY	2461	TTGGCGAAAGATCCACGTTATTTGGCGCCACCATGGAAAAACGCCAGAAAAGGTGAAATCATG	2520
Db	2521	TTGGCGAAAGATCCACGTTATTTGGCGCCACCATGGAAAAACGCCAGAAAAGGTGAAATCATG	2580
QY	2521	CCGAACATCCCGCAGATGTCGGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC	2580
Db	2581	CCGAACATCCCGCAGATGTCGGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC	2640
QY	2581	GCCAGCGTCTGACACTGTGATGAAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG	2640
Db	2641	GCCAGCGTCTGACACTGTGATGAAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG	2700
QY	2641	AACAAACAAACAATTAACAATAACAACTCGGGATCGAGGGAAAGGATTCAGAAATTC	2700
Db	2701	AACAAACAAACAATTAACAATAACAACTCGGGATCGAGGGAAAGGATTCAGAAATTC	2760
QY	2701	GGATCCTCT 2709	
Db	2761	GGATCCTCT 2769	

Search completed: August 1, 2005, 17:22:05
Job time : 9284 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2005, 20:22:50 ; Search time 9276 Seconds
(without alignments)
17238.272 Million cell updates/sec

Title: US-09-765-555B-14
Perfect score: 3300
Sequence: 1 ccgacacatcgatggtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hcg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_ats.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	3300	100.0	3300	6	AX202426 Sequence
2	3300	100.0	3300	6	AX202427 Sequence
3	3226.4	97.8	3300	6	AX202430 Sequence
4	3215.2	97.4	3300	6	AX202429 Sequence
5	3207.2	97.2	3300	6	AX202428 Sequence
6	2705.8	82.0	7373	6	AX284144 Sequence
7	2705.8	82.0	8101	6	AX172306 Sequence
8	2701	81.8	6648	6	AX377531 Sequence
9	2700	81.8	7475	6	AX91965 Sequence
10	2700	81.8	7475	6	AR031992 Sequence
11	2700	81.8	7475	6	AR207294 Sequence
12	2698	81.8	9191	6	AX377532 Sequence
13	2676.2	81.1	6724	12	AF097412 Sequence
14	2636.8	79.9	6806	6	AX378208 Sequence
15	2631	79.7	6706	12	AF031088 Sequence
16	2598.2	78.7	6748	12	AF031813 Sequence
17	1609.4	48.8	5558	6	CQ794769 Sequence
18	1474	47.0	4700	6	CQ846980 Sequence
19	1462.4	44.3	4921	6	AX768201 Sequence

20	1462.4	44.3	4935	6	AX768200	AX768200 Sequence
21	1462.4	44.3	4945	6	AX768202	AX768202 Sequence
22	1462.4	44.3	4951	6	AX768203	AX768203 Sequence
c	23	1432.4	43.4	6755	12	AF140577 Integrati
	24	1417.6	43.0	1511	6	AX431576 Sequence
	25	1417.6	43.0	4297	12	AY219685 Expressio
	26	1417.6	43.0	4359	12	AY219686 Expressio
	27	1417.6	43.0	4419	12	AY219688 Expressio
	28	1417.6	43.0	4855	12	AY219689 Expressio
	29	1417.6	43.0	4968	12	AY219687 Expressio
	30	1417.6	43.0	6956	12	AY219682 Expressio
	31	1417.6	43.0	7018	12	AY219683 Shuttle e
	32	1417.6	43.0	7509	12	AY219684 Shuttle e
	33	1400.8	42.4	4566	12	AF177933 Cloning v
	34	1400.8	42.4	4635	12	AF177932 Cloning v
	35	1400.8	42.4	6399	12	AY302759 Expressio
	36	1400.8	42.4	6587	12	AY289542 Expressio
	37	1362	41.3	5201	6	AR493834 Sequence
	38	1362	41.3	5201	6	AR493835 Sequence
	39	1329.2	40.3	4806	12	AF050464 Expressio
c	40	1238.4	37.5	4803	12	AY243506 Cloning v
c	41	1238.4	37.5	5024	6	CQ846977 Sequence
	42	1237.2	37.5	1922	6	AX478528 Sequence
	43	1204	36.5	4506	12	AB116367 Expressio
	44	1204	36.5	4557	6	E13740 DNA sequenc
	45	1204	36.5	4557	6	AR373251 Sequence

ALIGNMENTS

RESULT 1
AX202426
LOCUS AX202426 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 14 from Patent WO0152620.
ACCESSION AX202426
VERSION AX202426.1 GI:15392173
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 14 26-JUL-2001;
The Scripps Research Institute (US); SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES
source Location/Qualifiers
1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-ml and zinc finger protein 2Fpml"

ORIGIN	Query Match	100.0%;	Score 3300;	DB 6;	Length 3300;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 3300;	Conservative	0;	Mismatches	0;
				Indels	0;
				Gaps	0;
QY	1	CCGACCATCATGAATGGTGC	AAACCTTTCCGCGGTATG	CGCATGATGAGCCCGGAAGA	GA 60
Db	1	CCGACCATCATGAATGGTGC	AAACCTTTCCGCGGTATG	CGCATGATGAGCCCGGAAGA	GA 60
QY	61	GTCAATTCAGGTGGTGAATG	TGAACAGTAACTTATACG	TATCCGATGCGCAGATGCGG	120
Db	61	GTCAATTCAGGTGGTGAATG	TGAACAGTAACTTATACG	TATCCGATGCGCAGATGCGG	120
QY	121	GTGTCCTTTATCAGACCGG	TTTCCCGCGGTGGTGAAC	CCAGCCAGCCACCGTTTCTG	CGAAAA 180
Db	121	GTGTCCTTTATCAGACCGG	TTTCCCGCGGTGGTGAAC	CCAGCCAGCCACCGTTTCTG	CGAAAA 180
QY	181	CGCGGAAAAAGTGGAAGCG	CGCATGGCGGAGTGAATTA	TATCCCAACCGCGTGGCAC	240

Db	181	CGCGGAAAAAGTGGAAAGCGGCGATGGCGAGCTGAATTCACATTCACCAACCGCTGGCAC	240
Qy	241	AACAACTGGCGGGCAAAAGTCGTTGCTGATTTGGCGTGTGCCACCTCCAGTCTGGGCCCTTGC	300
Db	241	AACAACTGGCGGGCAAAAGTCGTTGCTGATTTGGCGTGTGCCACCTCCAGTCTGGGCCCTTGC	300
Qy	301	ACGCGCGTTCGCAAAATTTGTTCGGGGCGATTAATACTCGCGCCGATCAACTGGGTGCCACGC	360
Db	301	ACGCGCGTTCGCAAAATTTGTTCGGGGCGATTAATACTCGCGCCGATCAACTGGGTGCCACGC	360
Qy	361	TGGTGGTTCGATGGTAGAACAAGCGCGTCGGAAGCCTGTAAAGCGCGGTGCACAATC	420
Db	361	TGGTGGTTCGATGGTAGAACAAGCGCGTCGGAAGCCTGTAAAGCGCGGTGCACAATC	420
Qy	421	TTCTCGCGCAACGCGTCAAGTGGGTGATCAATTAACCTACCGCTGGATCACCAGGATGCCA	480
Db	421	TTCTCGCGCAACGCGTCAAGTGGGTGATCAATTAACCTACCGCTGGATCACCAGGATGCCA	480
Qy	481	TTGCTGTGGAAGCTCCCTGCACTAATGTTCCGGCGTTATTTCTTGATCTCTGACCAAGA	540
Db	481	TTGCTGTGGAAGCTCCCTGCACTAATGTTCCGGCGTTATTTCTTGATCTCTGACCAAGA	540
Qy	541	CACCCATCAACAGTATTATTTTCTCCCATGAAGACGTCACGCACTGGCGCTGGAGCATC	600
Db	541	CACCCATCAACAGTATTATTTTCTCCCATGAAGACGTCACGCACTGGCGCTGGAGCATC	600
Qy	601	TGGTCGCAATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAATGTCGTCTCGG	660
Db	601	TGGTCGCAATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAATGTCGTCTCGG	660
Qy	661	CGCGTCTCGCTCTGGCTGGCATTAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Db	661	CGCGTCTCGCTCTGGCTGGCATTAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Qy	721	CGAAACGGGAAGGCGACTGGAAGTGCATGTCGGTTTTCAACAAACCAATGCAATGCTGA	780
Db	721	CGAAACGGGAAGGCGACTGGAAGTGCATGTCGGTTTTCAACAAACCAATGCAATGCTGA	780
Qy	781	ATGAGGGCATCGTTCCCACTCGCATGCTGTGTCGAACGATCAGATGCGCTGGGCGCA	840
Db	781	ATGAGGGCATCGTTCCCACTCGCATGCTGTGTCGAACGATCAGATGCGCTGGGCGCA	840
Qy	841	TGCGCGCATTTACCGAGTTCGCGGTTCGCGCTTGGTGGCGATATCTCGTAGTGGGATACG	900
Db	841	TGCGCGCATTTACCGAGTTCGCGGTTCGCGCTTGGTGGCGATATCTCGTAGTGGGATACG	900
Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAACCAACCATCAACAGATTTTC	960
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCGCGCGTTAACCAACCATCAACAGATTTTC	960
Qy	961	GCCTGCTGGGGCAAAACAGCGTGAACCGCTTGTGTCGCAATCTCTCAGGGCCAGCGGTGA	1020
Db	961	GCCTGCTGGGGCAAAACAGCGTGAACCGCTTGTGTCGCAATCTCTCAGGGCCAGCGGTGA	1020
Qy	1021	AGGGCAATCAGCTGTCGCGCTCATCTGGTGAAGAAAGAAACCAACCTGGCGGCCAATA	1080
Db	1021	AGGGCAATCAGCTGTCGCGCTCATCTGGTGAAGAAAGAAACCAACCTGGCGGCCAATA	1080
Qy	1081	CGCAAAACCGCTCTCTCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCAGCAGGTTT	1140
Db	1081	CGCAAAACCGCTCTCTCCCGCGCGTTGGCCGATTCATTAATGCAGCTGGCAGCAGGTTT	1140
Qy	1141	CCGCACTGGAAGCGGCGAGTGACGGCAACGCAATTAATGTAGTTAGTCACTCATTTAG	1200
Db	1141	CCGCACTGGAAGCGGCGAGTGACGGCAACGCAATTAATGTAGTTAGTCACTCATTTAG	1200
Qy	1201	GCACAAATTCATGTTTACAGCTTTATCATCGACTGACGCGTGCACCAATGCTTCTCGCG	1260
Db	1201	GCACAAATTCATGTTTACAGCTTTATCATCGACTGACGCGTGCACCAATGCTTCTCGCG	1260
Qy	1261	TCAGGAGCCATCGGAAGCTGTGGTATGGCTGTGCGAGGTGTAATCACTGCATAATTCG	1320

Db	1261	TCAGCGACCATCGAAAGCTGTGGTATGGCTGTGTGACGGTTCGTAATCATCTGCATGATAATTCTG	1320
Qy	1321	TGTCGCTCAAGCGCACAATCCCGTCTCGATAAATGTTTTTTTGGCCGACATCATAAACGGTT	1380
Db	1321		
Qy	1381	CTGGCAAAATATCTCGAAATGAGCTGTTTGACAAATTAATCATCGGCTCGTAAATGTTGTGGA	1440
Db	1381		
Qy	1441	ATTGTGAGCGGATAACAATTTTCAACAGGAAAACAGCCAGTCCGTTTAGTGTGTTTTTCAACGA	1500
Db	1441		
Qy	1501	GCATCTTCAACCAACGAAGCACCATAGATTATGAAAATCTGAAGAAGTAAATCGGTAAATCTCG	1560
Db	1501	GCATCTTCAACCAACGAAGCACCATAGATTATGAAAATCTGAAGAAGTAAATCTGTGTAATCTCG	1560
Qy	1561	ATTAAACGGCGATAAAGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAGAAT	1620
Db	1561		
Qy	1621	ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCTGGAAGAGAAATCCACACAGGTT	1680
Db	1621	ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAATCTGGAAGAGAAATCCACACAGGTT	1680
Qy	1681	CGCGCAACTCGCGATGCGCTCGATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
Db	1681	CGCGCAACTCGCGATGCGCTCGATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
Qy	1741	GCTCAATCTGGCCTGTGGCTGAAATCACCCGGACAAAGCGTTCCAGCAACAGCTGTAT	1800
Db	1741		
Qy	1801	CCGTTTACCTGGGATGCGGTACCGCAAGCTGATTGCTTACCCGATCGCTGTT	1860
Db	1801	CCGTTTACCTGGGATGCGGTACCGCAAGCTGATTGCTTACCCGATCGCTGTT	1860
Qy	1861	GAGCGTTATGCTGATTATTAACAAGATCTCTGCGGAAACCCGCAAAAACCTGGGAA	1920
Db	1861	GAGCGTTATGCTGATTATTAACAAGATCTCTGCGGAAACCCGCAAAAACCTGGGAA	1920
Qy	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGGAAAGTAAAGCGCGCTGATGTTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGGAAAGTAAAGCGCGCTGATGTTTCAAC	1980
Qy	1981	CTGCAAGAACGGTACTTCACTGCGCGTGAATTGCTGCTGAACGGGGTTATGCGTTCAAG	2040
Db	1981	CTGCAAGAACGGTACTTCACTGCGCGTGAATTGCTGCTGAACGGGGTTATGCGTTCAAG	2040
Qy	2041	TATGAAAACGGCAAGTACGACATTAAGAAGCTGGGCTGGATAAACGTCGCGCGAAAGCG	2100
Db	2041	TATGAAAACGGCAAGTACGACATTAAGAAGCTGGGCTGGATAAACGTCGCGCGAAAGCG	2100
Qy	2101	GGTCTGACCTTCTGTTGACCTGATTAATAAACAACATGAATGCAGACACCGATTAC	2160
Db	2101	GGTCTGACCTTCTGTTGACCTGATTAATAAACAACATGAATGCAGACACCGATTAC	2160
Qy	2161	TCCATCGCAGAAGCTGCTTTTATTAAGCGGAAACAGCGATGACCATCAACGGCCGCTGG	2220
Db	2161	TCCATCGCAGAAGCTGCTTTTATTAAGCGGAAACAGCGATGACCATCAACGGCCGCTGG	2220
Qy	2221	GCATGFTCCAACATCGACACCGCAAAGTGAATTAATGTTGAACCGTACTGCGGACCTTC	2280
Db	2221	GCATGFTCCAACATCGACACCGCAAAGTGAATTAATGTTGAACCGTACTGCGGACCTTC	2280
Qy	2281	AAGGFTCAACCATCCAAACGGTTGCGTGTGAGCGGAGGATTAACCGCGCCAGT	2340
Db	2281	AAGGFTCAACCATCCAAACGGTTGCGTGTGAGCGGAGGATTAACCGCGCCAGT	2340
Qy	2341	CCGAAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACATCTGCTGACTGATGAAGGTCGTG	2400
Db	2341	CCGAAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACATCTGCTGACTGATGAAGGTCGTG	2400

QY	2401	GAAGCGGTAAATAAGACAAACCGCTGGGTGGCGGTAGCGCTGAAGCTTTACAGGAAGAG	2460
Db	2401		
QY	2461	TTGGCGAAAGATCCACGCTATTGCGGCCACCATGGAAGAACGCCAGAAAGGTGAATCATG	2520
Db	2461		
QY	2521	CCGAACATCCCGCAGATGTCGCTTTCTTGATGATCCGCTGATCAACGCC	2580
Db	2521		
QY	2581	GCCAGGGTCGTACAGCTGTGATGAAGCCCTGAAAGAGCGCCAGACTAAATTCGAGCTCG	2640
Db	2581		
QY	2641	AACAAACAAACAAATAAACAATAACCAACCTCGGATCGAGGGAAGGATTTCAGAATTC	2700
Db	2641		
QY	2701	GGATCCTCTTCTGTGSCCCAGGCGCCCTCGAGCCCGGGGAGAAGCCCTATGCTTGT	2760
Db	2701		
QY	2761	CCGGAATGTGTAAGTCTTCTCTCAGAGCTCTACCTGGTGGGCCACACAGGCTACCCAC	2820
Db	2761		
QY	2821	ACGGGTGAAAACCGTATAAATGCCAGAGTGGCGCAATCTTTAGCCAGTCCAGCAAC	2880
Db	2821		
QY	2881	CTGGTGGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGTCCAGAAATGTGGC	2940
Db	2881		
QY	2941	AAGTCTTTCTCGGTCTGACAAATCTGTCGGGCAACCAACGTACTCACACCGGGGAGAG	3000
Db	2941		
QY	3001	CCCTATGCTTGTCCGGAATGTGGTAAATGCCAGAGTGGCGCAATCTTTAGCCAGTCCAGCAAC	3060
Db	3001		
QY	3061	CAGCGTACCCACACGCGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGC	3120
Db	3061		
QY	3121	CAGGCGGGCCACTGGCCAGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT	3180
Db	3121		
QY	3181	CCAGAATGTGGCAAGTCTTCTCTCGGTCTGCAATCTCGTCCGCGACCAACGTA	3240
Db	3181		
QY	3241	ACCGGTAAAAAAGTAGTGGCCAGGCGCGCAGTACCGGTACGACGTTCCGGA	3300
Db	3241		

RESULT 2

AX202427

LOCUS

AX202427

DEFINITION

Sequence 15 from Patent WO0152620.

ACCESSION

AX202427

VERSION

AX202427.1

KEYWORDS

GI:15392175

SOURCE

synthetic construct

ORGANISM

synthetic construct

REFERENCE

1

AUTHORS	Barbas, C.F., Stege, J.T., Guan, X. and Dalmia, B.									
TITLE	Methods and compositions to modulate expression in plants									
JOURNAL	Patent: WO 0152620-A 15 26-JUL-2001; SYNGENTA AGRICULTURAL DISCOVERY, INC. (CA)									
FEATURES	Location/Qualifiers									
source	1..3300									
	/organism="synthetic construct"									
	/mol_type="unassigned DNA"									
	/db_xref="taxon:32630"									
	/notes="Partial sequence of pMal-m2 and zinc finger protein ZFPm2"									
ORIGIN										
Query Match	100.0%; Score 3300; DB 6; Length 3300;									
Best Local Similarity	100.0%; Pred. No. 0;									
Matches 3300; Conservative	0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	CCGACACCATCGAATGGTGC	AAACCTTTTCGCGGTATG	CGCATATGCGCATATAGCGCCCGG	MAAGAGA	60				
Db	1	CCGACACCATCGAATGGTGC	AAACCTTTTCGCGGTATG	CGCATATGCGCATATAGCGCCCGG	MAAGAGA	60				
QY	61	GTCAATTCAGGGTGGTGAAT	GTGAAACCACTGTAACGTT	TATACGATGTCGACAGATGATCCG		120				
Db	61	GTCAATTCAGGGTGGTGAAT	GTGAAACCACTGTAACGTT	TATACGATGTCGACAGATGATCCG		120				
QY	121	GTGTCCTTTATCAGACCG	TTTCCGCGGTGGTGAAC	CAGCGCAGCCAGCCAGCCAG	CGTTCCTGCG	180				
Db	121	GTGTCCTTTATCAGACCG	TTTCCGCGGTGGTGAAC	CAGCGCAGCCAGCCAGCCAG	CGTTCCTGCG	180				
QY	181	CGCGGGAAGTGAAGCGG	CGATGGCGGAGTGAAT	TACATTTCCCAACCGCGTGG	CAC	240				
Db	181	CGCGGGAAGTGAAGCGG	CGATGGCGGAGTGAAT	TACATTTCCCAACCGCGTGG	CAC	240				
QY	241	AACAACTGCGGGCAAC	ACAGTCTGCTGATTGCG	CGTTGGCCACTCCAGTCTGG	CCCTGC	300				
Db	241	AACAACTGCGGGCAAC	ACAGTCTGCTGATTGCG	CGTTGGCCACTCCAGTCTGG	CCCTGC	300				
QY	301	ACGCGCGCTCGCAAA	TGTGCGCGGATTAAT	TCTCGCGCGATCAACTGGGTGCC	ACG	360				
Db	301	ACGCGCGCTCGCAAA	TGTGCGCGGATTAAT	TCTCGCGCGATCAACTGGGTGCC	ACG	360				
QY	361	TGGTGTGTCGATGTG	TAGAACGAAGCGGCTCG	AAGCTGTAAAGCGGCGGTGC	ACAATC	420				
Db	361	TGGTGTGTCGATGTG	TAGAACGAAGCGGCTCG	AAGCTGTAAAGCGGCGGTGC	ACAATC	420				
QY	421	TTCTCGGCAACGCGTCA	GTTGGCTGATCAATTA	ACTATCCGCTGGATGAC	CGAGTGCCA	480				
Db	421	TTCTCGGCAACGCGTCA	GTTGGCTGATCAATTA	ACTATCCGCTGGATGAC	CGAGTGCCA	480				
QY	481	TTGCTGTGAAGCTGCTG	CATTAATGTTCCGCGGT	TATTTCTTGATGTCTCTG	ACCAGA	540				
Db	481	TTGCTGTGAAGCTGCTG	CATTAATGTTCCGCGGT	TATTTCTTGATGTCTCTG	ACCAGA	540				
QY	541	CACCCATCAACAGTAT	TATTTCTCCATGAAG	ACGGTACGCGACTGGGCGTGG	AGCATC	600				
Db	541	CACCCATCAACAGTAT	TATTTCTCCATGAAG	ACGGTACGCGACTGGGCGTGG	AGCATC	600				
QY	601	TGGTGCATTTGGGTCA	CCAGCAATCGCGCTGT	TAGCGGGCCCATTAAGTCTGT	CTCTCGG	660				
Db	601	TGGTGCATTTGGGTCA	CCAGCAATCGCGCTGT	TAGCGGGCCCATTAAGTCTGT	CTCTCGG	660				
QY	661	CGCGTCTGCGCTCTG	GGCTGGCATAAATATCT	CACTCGCAATCAAAATTCAG	CCGATAG	720				
Db	661	CGCGTCTGCGCTCTG	GGCTGGCATAAATATCT	CACTCGCAATCAAAATTCAG	CCGATAG	720				
QY	721	CGGAACGGGAAGCGG	ACTGGAGTGCCATGTC	CGGTTTTTCAACAAACCAT	GCAAAATGCTGA	780				
Db	721	CGGAACGGGAAGCGG	ACTGGAGTGCCATGTC	CGGTTTTTCAACAAACCAT	GCAAAATGCTGA	780				
QY	781	ATGAGGCGATCTGTT	CCCACTGCGATGCTG	TGTCGAACGATCAGATGG	CGCTGGGCGCAA	840				
Db	781	ATGAGGCGATCTGTT	CCCACTGCGATGCTG	TGTCGAACGATCAGATGG	CGCTGGGCGCAA	840				

841 TCGCGCCATTACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGTAGTGGGATACG 900
Db
841 TCGCGCCATTACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTATATATCCCGCGTTAAACCAATCAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTATATATCCCGCGTTAAACCAATCAACAGGATTTTC 960
Qy 961 GCGTGTGGGGCAAAACAGCGTGGACGCTTCTGCACTCTCTCAGGCGCACGGGTGA 1020
Db 961 GCGTGTGGGGCAAAACAGCGTGGACGCTTCTGCACTCTCTCAGGCGCACGGGTGA 1020
Qy 1021 AGGCAATCAGCTGTGTCCTGCTCACTGGTGAAGAAACCAACCTGCGGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGTCCTGCTCACTGGTGAAGAAACCAACCTGCGGCCCAATA 1080
Qy 1081 CGCAACCGCCTCTCCCGCGGTTGGCGATTCATTAATGAGCTGGCAACAGGTTT 1140
Db 1081 CGCAACCGCCTCTCCCGCGGTTGGCGATTCATTAATGAGCTGGCAACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGCAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGCAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTAG 1200
Qy 1201 GCACAATTTCTATGTTGACGCTTATCATGACTGCAGGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTTCTATGTTGACGCTTATCATGACTGCAGGTGCACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGCGCATCGAAGCTGTGGTATGCTGTGAGCTGCAAGTGCATTAATTCG 1320
Db 1261 TCAGGCGCATCGAAGCTGTGGTATGCTGTGAGCTGCAAGTGCATTAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGCGACATCATACGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGCGACATCATACGTT 1380
Qy 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
Qy 1441 ATTGTAGCGGATACAAATTTACAGAGAAACAGCCAGTCCGTTAGTGTTCACGA 1500
Db 1441 ATTGTAGCGGATACAAATTTACAGAGAAACAGCCAGTCCGTTAGTGTTCACGA 1500
Qy 1501 GCACCTTCAACAAAGGACCATAGATTATGAAAACCTGAAGAGTAAACTGGTAACTCGG 1560
Db 1501 GCACCTTCAACAAAGGACCATAGATTATGAAAACCTGAAGAGTAAACTGGTAACTCGG 1560
Qy 1561 ATTAAACGGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAAACGGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAACTGGAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAACTGGAGAGAAATTCACAGGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTTCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTTCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGCTGAATATCAACCGGACAAAGCGTTCCAGGAACAGCTGAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAATATCAACCGGACAAAGCGTTCCAGGAACAGCTGAT 1800
Qy 1801 CCGTTTACTGGGATGCGGTACGTTTACACGCAAGCTGATTTGCTTACCGATCGCTGTT 1860
Db 1801 CCGTTTACTGGGATGCGGTACGTTTACACGCAAGCTGATTTGCTTACCGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920

1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTATCTTCACTGCGCGCTGATGCTGCTGACGCGGGTATTCGGTTCAG 2040
Db 1981 CTGCAAGAACCGTATCTTCACTGCGCGCTGATGCTGCTGACGCGGGTATTCGGTTCAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGCAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGCAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGTGTGACCTGATTAATAAACAACAACATGATGACACCGATTC 2160
Db 2101 GGTCTGACCTTCTCGTGTGACCTGATTAATAAACAACAACATGATGACACCGATTC 2160
Qy 2161 TCCATCGAAGAGTGGCTTTAATAAGGCGAAACAGCGATGACATCAACGGCGCGTGG 2220
Db 2161 TCCATCGAAGAGTGGCTTTAATAAGGCGAAACAGCGATGACATCAACGGCGCGTGG 2220
Qy 2221 GCATGGTCCAAACATCGACACCGCAAGAGTGAATTTATGGTGAACGGTACTGCCACCTTC 2280
Db 2221 GCATGGTCCAAACATCGACACCGCAAGAGTGAATTTATGGTGAACGGTACTGCCACCTTC 2280
Qy 2281 AAGGCTCAACCATCCAAACCGTTCGTTGGCGTGTGAGCGCAGGTATTAACCGCGCCAGT 2340
Db 2281 AAGGCTCAACCATCCAAACCGTTCGTTGGCGTGTGAGCGCAGGTATTAACCGCGCCAGT 2340
Qy 2341 CCGAAACAAAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGCTG 2400
Db 2341 CCGAAACAAAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGCTG 2400
Qy 2401 GAACGGTTAATAAGACAAACCGTGGTGGCTAGCGTGAAGTCTTACGAGGAGAG 2460
Db 2401 GAACGGTTAATAAGACAAACCGTGGTGGCTAGCGTGAAGTCTTACGAGGAGAG 2460
Qy 2461 TTGCGAAAGATCCACGCTATTGCGCCCAACCATGAAAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGCGAAAGATCCACGCTATTGCGCCCAACCATGAAAAACGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGGCTTCTGATGTCGGTGGTACTGCGGTGATCAAGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGGCTTCTGATGTCGGTGGTACTGCGGTGATCAAGCC 2580
Qy 2581 GCCAGCGTCTGACAGTGTGATGAAGCCTGAAAGCGCGCAGACTAATTCAGGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTGTGATGAAGCCTGAAAGCGCGCAGACTAATTCAGGCTCG 2640
Qy 2641 AACCAACAAACATTAACAACAACCTCGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACCAACAAACATTAACAACAACCTCGGATCGAGGGAAGGATTCAGAAATTC 2700
Qy 2701 GGATCTCTTCTCTGTTGGCCAGCGGCTCTGAGCCCGGGGAGAGCCCTATGCTGT 2760
Db 2701 GGATCTCTTCTCTGTTGGCCAGCGGCTCTGAGCCCGGGGAGAGCCCTATGCTGT 2760
Qy 2761 CCGGAATGTGTAAGTCTCTCTCAGAGTCTCACTGGTGGCCACAGCGTACCCAC 2820
Db 2761 CCGGAATGTGTAAGTCTCTCTCAGAGTCTCACTGGTGGCCACAGCGTACCCAC 2820
Qy 2821 ACGGTTGAAAAACCGTATAAATGCCAGAGTGGCGAAAATCTTTTACCGAGTCCAGCAAC 2880
Db 2821 ACGGTTGAAAAACCGTATAAATGCCAGAGTGGCGAAAATCTTTTACCGAGTCCAGCAAC 2880
Qy 2881 CTGGTGGCCATCAACGCTCTACTTGGGAGAGCCATACAAATCTCCAGATGTGGC 2940
Db 2881 CTGGTGGCCATCAACGCTCTACTTGGGAGAGCCATACAAATCTCCAGATGTGGC 2940
Qy 2941 AAGTCTTCTCTCGGTGCAATCTCGTCCGCAACCAAGTACTCACACCGGAGAGAG 3000
Db 2941 AAGTCTTCTCTCGGTGCAATCTCGTCCGCAACCAAGTACTCACACCGGAGAGAG 3000
Qy 3001 CCCTATGCTTGTCCGGAAATGTGTAAGTCTTTCAGCGCAGGATAACTGCTGGTGGCCAC 3060

Db 3001 ||||| CCCTATGCTTGTCCGGAATGTGGTAAGTCTTTACGCGCAGCGATAAACCCTGTGTGCGCCAC 3060
Qy 3061 CAGCGGTACCCACACAGGGGTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTTAGC 3120
Db 3061 CAGCGGTACCCACACAGGGGTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTTAGC 3120
Qy 3121 CAGGCGGCGCACTCGGCCAGCCATCAAGCAGCTCATCTACTGCGGAGAGCATTACAAATGT 3180
Db 3121 CAGGCGGCGCACTCGGCCAGCCATCAAGCAGCTCATCTACTGCGGAGAGCATTACAAATGT 3180
Qy 3181 CCAGAATGTGGCAAGCTTTCTCTCGGTCTGACATCTGCTCGGCACCAAGTACTCAC 3240
Db 3181 CCAGAATGTGGCAAGCTTTCTCTCGGTCTGACATCTGCTCGGCACCAAGTACTCAC 3240
Qy 3241 ACCGGTAAAAAATAGTGGCCAGGCGCGCCAGTACCCCGTACGAGTTCGCGACTACGCT 3300
Db 3241 ACCGGTAAAAAATAGTGGCCAGGCGCGCCAGTACCCCGTACGAGTTCGCGACTACGCT 3300

RESULT 3
AX202430
LOCUS AX202430 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 18 from Patent WO0152620.
ACCESSION AX202430
VERSION AX202430.1 GI:15392178
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
TITLE Methods and compositions to modulate expression in plants
JOURNAL Patent: WO 0152620-A 18 26-JUL-2001;
The Scripps Research Institute (US); SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)

FEATURES
source
1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Partial sequence of pMal-Ap3 and zinc finger protein ZFPp3"

ORIGIN
Query Match 97.8%; Score 3226.4; DB 6; Length 3300;
Best Local Similarity 98.6%; Pred.No. 0; Mismatches 0; Indels 0; Gaps 0;
Matches 3254; Conservative 0;

Qy 1 CGGACACCATCGAATGGTGC AAAACCTTTTCGGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 1 CGGACACCATCGAATGGTGC AAAACCTTTTCGGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
Qy 61 GTCAATTACGGGTGGTGAATGTGA AACACAGTAACGTTATACGATGTGCGAGATATGCCG 120
Db 61 GTCAATTACGGGTGGTGAATGTGA AACACAGTAACGTTATACGATGTGCGAGATATGCCG 120
Qy 121 GTGTCTCTTATCAGACCGTTTC CGCGGTGTGAAACAGGCGCCAGCCAGCTTTCTGCGAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTC CGCGGTGTGAAACAGGCGCCAGCCAGCTTTCTGCGAAA 180
Qy 181 CGCGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
Qy 241 AACAACTGCGGGCAAAAGTGTG TGTCTGATTGGCGTTCGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGCGGGCAAAAGTGTG TGTCTGATTGGCGTTCGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGCCCGTTCGCAATTTGTG CGGGCGATTAATCTCGGCCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCCCGTTCGCAATTTGTG CGGGCGATTAATCTCGGCCGATCAACTGGGTGCCAGCG 360

Qy 361 TGGTGGTTCGATGGTGAAGCGAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCACAATC 420
Db 361 TGGTGGTTCGATGGTGAAGCGAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCACAATC 420
Qy 421 TTCTCGCGCAACGGGTGAGTGGGCTGATCAATTAATCTATCCGCTGATCAACGAGATGCCA 480
Db 421 TTCTCGCGCAACGGGTGAGTGGGCTGATCAATTAATCTATCCGCTGATCAACGAGATGCCA 480
Qy 481 TTGCTGTGGAAGCTCGCTTGCATTAATGTTCCGGCGCTTATTTCTTGATGTCCTGACCGA 540
Db 481 TTGCTGTGGAAGCTCGCTTGCATTAATGTTCCGGCGCTTATTTCTTGATGTCCTGACCGA 540
Qy 541 CACCCATCAACAGTATTAATTTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC 600
Qy 601 TGCTCGCATTCGGTCCACGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db 601 TGCTCGCATTCGGTCCACGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Qy 661 CGCGTCTCGCTCGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTCGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATGTTCCCACTCGCATGCTGTTGCAACAGATCAGATGCGCTGGCGGCA 840
Db 781 ATGAGGGCATGTTCCCACTCGCATGCTGTTGCAACAGATCAGATGCGCTGGCGGCA 840
Qy 841 TCGCGGCCATTAACGAGTCCGGGCTCGCGTGTGGTGGCGATATCTCGGTAGTGGGATAG 900
Db 841 TCGCGGCCATTAACGAGTCCGGGCTCGCGTGTGGTGGCGATATCTCGGTAGTGGGATAG 900
Qy 901 ACGATACCGAAGACAGTCTATGTTATATCCCGCTTAAACCAATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGTCTATGTTATATCCCGCTTAAACCAATCAAAACAGGATTTTC 960
Qy 961 GCCTGCTGGGCAACACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGTGA 1020
Qy 1021 AGGCAATCAGCTGTGCGCGCTCTCACTGGTGAAAGAAAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGCGCGCTCTCACTGGTGAAAGAAAAACCAACCTGGCGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGCGCTGGCAGCAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGCGCTGGCAGCAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTCACTATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTCACTATTAG 1200
Qy 1201 GCACAATTTCTCATGTTTGACAGCTTATCATGCACTGCAAGTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTTCTCATGTTTGACAGCTTATCATGCACTGCAAGTGCACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGACGCCATTCGGAAGCTGTGGTATGCGTGTGAGGTCGTAATCACTGCATTAATTCG 1320
Db 1261 TCAGGACGCCATTCGGAAGCTGTGGTATGCGTGTGAGGTCGTAATCACTGCATTAATTCG 1320
Qy 1321 TGTGCTCAAGCGGCACTCCCGTTCGGAATAATGTTTTTGGCGCGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCGGAATAATGTTTTTGGCGCGACATCATACGGTT 1380
Qy 1381 CTGGCAAAATTTCTGAAATGAGCTGTGCACAATTAATCATCGGCTCGTATATGTGTGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTGCACAATTAATCATCGGCTCGTATATGTGTGA 1440
Qy 1441 ATTGTGAGCGGATAACAAATTTTCAACAGGAAAAAGCCAGTCCGTTTAGGTGTTTTCACGA 1500

ORIGIN	ZFPm4"									
	Query Match	97.4%;	Score 3215.2;	DB 6;	Length 3300;					
	Best Local Similarity	98.4%;	Pred. No. 0;							
	Matches 3247;	Conservative	0;	Mismatches	53;	Indels	0;	Gaps	0;	
Qy	1	CCGACACCATCGAATGCTGCAAAACCTTTCCGCGTATGGCATGATAGCGCCCGGAAGAGA	60							
Db	1	CCGACACCATCGAATGCTGCAAAACCTTTCCGCGTATGGCATGATAGCGCCCGGAAGAGA	60							
Qy	61	GTCAAATTCAGGGTGGTGAATGTGAACCAAGTAAACGTTTATACGATGTCGCAGAGTATCCCG	120							
Db	61	GTCAAATTCAGGGTGGTGAATGTGAACCAAGTAAACGTTTATACGATGTCGCAGAGTATCCCG	120							
Qy	121	GTGTCCTTATACAGACCGTTTCCCGCGTGGTGAACCAAGCGCAGCCACGTTTCTGCGAAAA	180							
Db	121	GTGTCCTTATACAGACCGTTTCCCGCGTGGTGAACCAAGCGCAGCCACGTTTCTGCGAAAA	180							
Qy	181	CGCGGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAAATTAATTTCCCAACCGCGTGGCAC	240							
Db	181	CGCGGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAAATTAATTTCCCAACCGCGTGGCAC	240							
Qy	241	AACAACTGGCGGGCAAAACAGTCGTTGCTGATGTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300							
Db	241	AACAACTGGCGGGCAAAACAGTCGTTGCTGATGTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300							
Qy	301	ACGCGCGTGCACAAATGTCCGCGCGATTTAAATCTCCGCGCGATCAACTGGTGCCACGG	360							
Db	301	ACGCGCGTGCACAAATGTCCGCGCGATTTAAATCTCCGCGCGATCAACTGGTGCCACGG	360							
Qy	361	TGTTGGTGTGATGTAGAACGAAGCGCGCTCGAAGCCTGTAAAGCGCGCTGCACAATC	420							
Db	421	TTCTTCGCGCAACCGGTCACTGATGGGTGATCAATTAATCTATCGCTGGATGATCCAGATGCCA	480							
Qy	481	TTGCTGTGGAAAGCTCGCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCAGA	540							
Db	481	TTGCTGTGGAAAGCTCGCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTTCTCTGACCAGA	540							
Qy	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600							
Db	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600							
Qy	601	TGTTGCGCATTTGGGTCAACAGCAAAATCGCGTGTAGCGGGGCCAATTAAGTTCTGCTCTCGG	660							
Db	601	TGTTGCGCATTTGGGTCAACAGCAAAATCGCGTGTAGCGGGGCCAATTAAGTTCTGCTCTCGG	660							
Qy	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCCGATAG	720							
Db	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCCGATAG	720							
Qy	721	CGGAAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTTCAACAAACCATGCAAAATGCTGA	780							
Db	721	CGGAAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTTCAACAAACCATGCAAAATGCTGA	780							
Qy	781	ATGAGGGCATCGTTCCCACTCCGATGCTGGTTGTCGAACAGATCAGATGGCGCTGGCGCAA	840							
Db	781	ATGAGGGCATCGTTCCCACTCCGATGCTGGTTGTCGAACAGATCAGATGGCGCTGGCGCAA	840							
Qy	841	TGCGGGCCATTACCGAGTCCGGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATACG	900							
Db	841	TGCGGGCCATTACCGAGTCCGGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATACG	900							
Qy	901	ACGATACCGAAGACAGCTCATGTTTATATCCCGCGTTTAAACCAACCATCAAAACAGGATTTTC	960							
Db	901	ACGATACCGAAGACAGCTCATGTTTATATCCCGCGTTTAAACCAACCATCAAAACAGGATTTTC	960							
Qy	961	GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGCGTGA	1020							
Db	961	GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGCGTGA	1020							

```
QY 2101 GGTCTGACCTTCTGTGTTGACCTGATTAAAAACAAACACATGAATGCGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGTGTTGACCTGATTAAAAACAAACACATGAATGCGACACCGATTAC 2160
QY 2161 TCATTCGAGAGCTGCTCTTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCATTCGAGAGCTGCTCTTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGTCACACATCGACACGACGAAGTGAATTTATGGTGAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGTCACACATCGACACGACGAAGTGAATTTATGGTGAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCAAAACCGTTGTCGTCGCTGAGCGAGGTATTAAAGCCGCCAGT 2340
Db 2281 AAGGGTCAACCATCAAAACCGTTGTCGTCGCTGAGCGAGGTATTAAAGCCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTCACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTCACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGCTGCTGAGCGAGGTATTAAAGCCGCCAGT 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGCTGCTGAGCGAGGTATTAAAGCCGCCAGT 2460
QY 2461 TTGGCGAAGATCCACGTATTGCGGCCACCATGGAAGAGCGCCAGAAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGTATTGCGGCCACCATGGAAGAGCGCCAGAAAAGGTGAATCATG 2520
QY 2521 CCGAATCCCGCAGATGTCGCTTCTGTTATGCGGTGCGTACGCGGTGATCAAGCC 2580
Db 2521 CCGAATCCCGCAGATGTCGCTTCTGTTATGCGGTGCGTACGCGGTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGGAAGACGCGACACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGGAAGACGCGACACTAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTTCAGAAATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTTCAGAAATTC 2700
QY 2701 GGATCCTCTTCTCTGTGGCCAGCGGCCCTCGAGCCCGGGGAGAACCTATGCTTGT 2760
Db 2701 GGATCCTCTTCTCTGTGGCCAGCGGCCCTCGAGCCCGGGGAGAACCTATGCTTGT 2760
QY 2761 CCGGAATGTGGTAAGTCTCTCTCAGAGCTCTCACCTGGTGGCCACACGCGTACCCAC 2820
Db 2761 CCGGAATGTGGTAAGTCTCTCAGCCAGAGAGCTCCCTGGTGGCCACACGCGTACCCAC 2820
QY 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAATCTTTTAGCCAGTCCAGCAAC 2880
Db 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAATCTTTTAGCCAGAGCAGCAGC 2880
QY 2881 CTGGTGGCCCATCAACGCACTCATACTGGCAGAGGCCATCAAAATGTCAGAAATGGC 2940
Db 2881 CTGGTGGCCCATCAACGCACTCATACTGGCAGAGGCCATCAAAATGTCAGAAATGGC 2940
QY 2941 AAGTCTTTCTCTCGGTGTGACAATCTCGTCGGCACCAACGTACTCACACCGGGGAGAAG 3000
Db 2941 AAGTCTTTCTCTCGGTGTGACAATCTCGTCGGCACCAACGTACTCACACCGGGGAGAAG 3000
QY 3001 CCCTATGTTTCCGGAATGGTAAGTCTTTCAGCCGCGAGATTAACCTGGTGGCCAC 3060
Db 3001 CCCTATGTTTCCGGAATGGTAAGTCTTTCAGCCGCGAGATTAACCTGGTGGCCAC 3060
QY 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAATCTTTTAGC 3120
QY 3121 CAGCGCGGCCACCTGGCCAGCCATCAAGCACTCATACTGGCGGAGAGCCATCAAAATGT 3180
Db 3121 CAGCGCGGATAACCTGGTGGCGCATCAAGCACTCATACTGGCGGAGAGCCATCAAAATGT 3180
```

```
QY 3181 CCAGAAATGTGCAAGTCTTTCTCTCGGTCTGACAAATCTGTCGGCACCAACGTACTCAC 3240
Db 3181 CCAGAAATGTGCAAGTCTTTCTCTCGGTCTGACAAATCTGTCGGCACCAACGTACTCAC 3240
QY 3241 ACCGGTAAAAAAAATAGTGGCCAGCGCCAGTACCGTACGACGTTCCGGACTACGCT 3300
Db 3241 ACCGGTAAAAAAAATAGTGGCCAGCGCCAGTACCGTACGACGTTCCGGACTACGCT 3300

RESULT 5
AX202428
LOCUS AX202428 3300 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 16 from Patent WO0152620.
ACCESSION AX202428
VERSION AX202428.1 GI:15392176
KEYWORDS
SOURCE
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE
1 Barbas,C.F., Stege,J.T., Guan,X. and Dalmia,B.
AUTHORS Methods and compositions to modulate expression in plants
TITLE Patent: WO 0152620-A 16 26-JUL-2001;
JOURNAL The Scripps Research Institute (US) ; SYNGENTA AGRICULTURAL
DISCOVERY, INC. (CA)
FEATURES
Location/Qualifiers
source 1..3300
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/notes="Partial sequence of pMal-m3 and zinc finger protein ZFPm3"

ORIGIN
Query Match 97.2%; Score 3207.2; DB 6; Length 3300;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 3242; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 1 CCGACACCATCGATGATGTCGCAAAACCTTTCGGGTATGGCATATAGCGCCCGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTCGGGTATGGCATATAGCGCCCGAAGAGA 60
QY 61 GTCAATTCAGGCTGGTGAATGTGAACCACTGTAACGTTATACGATGTGCGAGAGTATGCCG 120
Db 61 GTCAATTCAGGCTGGTGAATGTGAACCACTGTAACGTTATACGATGTGCGAGAGTATGCCG 120
QY 121 GTGTCTCTTATACAGACCGTTTCCGGGTGGTGAACAGGCCACCGACCGTTTCTGCGAAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCGGGTGGTGAACAGGCCACCGACCGTTTCTGCGAAAA 180
QY 181 CCGGGGAAAAAGTGGAAAGCGCGATGCGGAGCTGAATTACATTCCCAACCGCGTGGCAC 240
Db 181 CCGGGGAAAAAGTGGAAAGCGCGATGCGGAGCTGAATTACATTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGCAACAGTCTGTTGCTGATTGGCGTTGGCCACTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAACAGTCTGTTGCTGATTGGCGTTGGCCACTCCAGTCTGGCCCTGC 300
QY 301 ACGGCGCTCGCAAAATGTCGGCGCATTAATATCTCGCGCCGATCAACTGGGTGCGACGC 360
Db 301 ACGGCGCTCGCAAAATGTCGGCGCATTAATATCTCGCGCCGATCAACTGGGTGCGACGC 360
QY 361 TGTGTGTGTCGATGTTAGAACGAAGCGCGTCCAAACCTGTAAAGCGCGGTGCACAAATC 420
Db 361 TGTGTGTGTCGATGTTAGAACGAAGCGCGTCCAAACCTGTAAAGCGCGGTGCACAAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAACATATCCGCTGGATGACCAAGTGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAACATATCCGCTGGATGACCAAGTGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAAATGTTCCGCGTTATTTCTTGATGTCTGTGACAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAAATGTTCCGCGTTATTTCTTGATGTCTGTGACAGA 540
```

Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGAGCATC 600
Qy 601 TGGTCGCAATCGGTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG 660
Db 601 TGGTCGCAATTTGGTCACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG 660
Qy 661 CGCGTCTCGCTCGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTCGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGACATGGAGTGCATGTCGGTTTCAACAAACCATGCAATGCTGA 780
Db 721 CGGAACGGGAAGGACATGGAGTGCATGTCGGTTTCAACAAACCATGCAATGCTGA 780
Qy 781 ATGAGGGCATGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGCGCTGGCGCAA 840
Db 781 ATGAGGGCATGTTCCCACTGCGATGCTGTTGCCAACGATCAGATGCGCTGGCGCAA 840
Qy 841 TCGCGCCATTACCGAGTCCGGGCTCGCGTGTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TCGCGCCATTACCGAGTCCGGGCTCGCGTGTGGTGGGATATCTCGGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACCAATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACCAATCAAAACAGGATTTTC 960
Qy 961 GCCTGCTGGGCAACACGAGCTGACCGCTGCTGCAACTCTCTCAGGCGCAGGGGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGCTGACCGCTGCTGCAACTCTCTCAGGCGCAGGGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAACCAACCTGGCGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGGACGACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGGACGACAGGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
Qy 1201 GCACAATTCATGTTTGACAGCTTATCATCGACTGCACGGTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTCATGTTTGACAGCTTATCATCGACTGCACGGTGACCAATGCTTCTGGCG 1260
Qy 1261 TCAGCAGCCATCGGAAGCTGGGTATGGCTGTGCAGGTCGTPAATCACTGCATAAATTCG 1320
Db 1261 TCAGCAGCCATCGGAAGCTGGGTATGGCTGTGCAGGTCGTPAATCACTGCATAAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGGACATCATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGGACATCATACGGTT 1380
Qy 1381 CTGGCAATATTTCTGAAATGAGCTGTGACAATTAATCATCGGCTCGTAAATGTTGGA 1440
Db 1381 CTGGCAATATTTCTGAAATGAGCTGTGACAATTAATCATCGGCTCGTAAATGTTGGA 1440
Qy 1441 ATTGTAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTTTTCAG 1500
Db 1441 ATTGTAGCGGATAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTTTTCAG 1500
Qy 1501 GCATTTACCAACAGGACCATAGATTATGAAAACTGAAGAGGTAAACCTGTAATCTGG 1560
Db 1501 GCATTTACCAACAGGACCATAGATTATGAAAACTGAAGAGGTAAACCTGTAATCTGG 1560
Qy 1561 ATTAACGGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTAACGGGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620

Qy 1621 ACCGGAATTAAGTCACCGTTGACATCCGGATAAACTGGAAGAGAAATTTCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCACCGTTGACATCCGGATAAACTGGAAGAGAAATTTCCACAGGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCGGAACAAAGCGTTCCAGGAACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCGGAACAAAGCGTTCCAGGAACAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGTACGTTCAACGGCAAGCTGATGTTCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGTACGTTCAACGGCAAGCTGATGTTCTTACCCGATCGCTGTT 1860
Qy 1861 GAAACGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAAACGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCGCTGGATFAAGAACTGAAACGGAAGAAAGCTGAGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATFAAGAACTGAAACGGAAGAAAGCTGAGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATGCTGCTGACGGGGTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATGCTGCTGACGGGGTATGCGTTCAAG 2040
Qy 2041 TATGAAACCGGCAAGTAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGGCAAGTAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGCTGCTGCTGATTTAAAAAACAACATGAAATGACAGACCCGATTTAC 2160
Db 2101 GGTCTGACCTTCTGCTGCTGCTGATTTAAAAAACAACATGAAATGACAGACCCGATTTAC 2160
Qy 2161 TCCATCCGAGAGCTGCTTTTAAAGGCGAAACAGAGTACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCCGAGAGCTGCTTTTAAAGGCGAAACAGAGTACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCCAACTCGACACGACCAAGTGAATTAATGGTGAACGCTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAACTCGACACGACCAAGTGAATTAATGGTGAACGCTACTGCGGACCTTC 2280
Qy 2281 AAGGGTCAACCATCCAAACCGTTGCTGGCGTGTGAGCGAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTGCTGGCGTGTGAGCGAGGTATTAACGCCGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGCTCG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGCTCG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTGAGGAGGTGAAATCATG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTGAGGAGGTGAAATCATG 2460
Qy 2461 TTGGCGAAAGATTCACGTTATTCGCGCCACCATGGAAGAAACGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGCGAAAGATTCACGTTATTCGCGCCACCATGGAAGAAACGCCAGAAAGGTGAAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGGTCGCTGCTGATGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCGGTCGCTGCTGATGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Qy 2641 AACAAACAAACAATAAACAACCAACCTTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACAAACAAACAATAAACAACCAACCTTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Qy 2701 GGATCCTCTTCTGCTGGCCAGCGGCCCTTCGAGACCGCGGGGAGAACCCCTATGCTGT 2760

841 TGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db TGCGGCCATTACCGAGTCCGGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCTTAACACCACTCAAAACAGGATTTTC 960
Db ACGATACCGAAGACAGCTCATGTTATATCCGCGCTTAACACCACTCAAAACAGGATTTTC 960
961 GCCTGCTGGGCAAAACCGAGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db GCCTGCTGGGCAAAACCGAGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
1021 AGGGAATCAGCTGTGTCGCTCTCACTGCTGGTGAACCAACCTCTCAGGGCCAGGCGGTGA 1080
Db AGGGAATCAGCTGTGTCGCTCTCACTGCTGGTGAACCAACCTCTCAGGGCCAGGCGGTGA 1080
1081 CGCAAAACCGCTCTCCCGCGGTTGGCGATTCAATTAATGCAAGTGGACGACAGGTTT 1140
Db CGCAAAACCGCTCTCCCGCGGTTGGCGATTCAATTAATGCAAGTGGACGACAGGTTT 1140
1141 CCCGACTGGAAGCGGCGAGTGGACGCAACGCAATTAATGAGTTAGTCACTCATTTAG 1200
Db CCCGACTGGAAGCGGCGAGTGGACGCAACGCAATTAATGAGTTAGTCACTCATTTAG 1200
1201 GCACAATTTCTCATGTTTGACAGCTTATCATGCACTGCAAGTGGACCAATGCTCTGCGG 1260
Db GCACAATTTCTCATGTTTGACAGCTTATCATGCACTGCAAGTGGACCAATGCTCTGCGG 1260
1261 TCAGGACGCAATCGGAAGCTGGGTATGGCTGTCAGGTGTAATCACTGCATTAATTCG 1320
Db TCAGGACGCAATCGGAAGCTGGGTATGGCTGTCAGGTGTAATCACTGCATTAATTCG 1320
1321 TGTGCTCAAGCGCACTCCCGTTCGGAATATGTTTTGGCGCGACATCATACGGTT 1380
Db TGTGCTCAAGCGCACTCCCGTTCGGAATATGTTTTGGCGCGACATCATACGGTT 1380
1381 CTGGCAAAATTTCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTAAATGTTGTTGA 1440
Db CTGGCAAAATTTCTGAAATGAGCTGTTGCAATTAATCATCGGCTCGTAAATGTTGTTGA 1440
1441 ATTGTGCGGATACAAATTTACACAGGAACCGCAGTCCGTTAGTGTGTTTCAGGA 1500
Db ATTGTGCGGATACAAATTTACACAGGAACCGCAGTCCGTTAGTGTGTTTCAGGA 1500
1501 GCACCTTCAACCAAGGACCATAGATTATGAAACTGGAAGAGTAAACTGGTAATCTGG 1560
Db GCACCTTCAACCAAGGACCATAGATTATGAAATCGAAGAGTAAACTGGTAATCTGG 1560
1561 ATTAAACGCGATPAAAGGCTATAACGGTCTCGCTGAAGTGGTAAGAAATTCGAGAAAGAT 1620
Db ATTAAACGCGATPAAAGGCTATAACGGTCTCGCTGAAGTGGTAAGAAATTCGAGAAAGAT 1620
1621 ACCGGAATTAAGTCAACGGTTGAGCATCGGATAAATCGGAAGAGAAATTCACAGGTT 1680
Db ACCGGAATTAAGTCAACGGTTGAGCATCGGATAAATCGGAAGAGAAATTCACAGGTT 1680
1681 GCGCAACTGGCGATGGCCCTGACATTATCTTCGGCACACACCGCTTTGGTGGCTAC 1740
Db GCGCAACTGGCGATGGCCCTGACATTATCTTCGGCACACACCGCTTTGGTGGCTAC 1740
1741 GCTCAATCTGGCCTGTTGGCTGAATATCAACCGGACAAAGCGTTCCAGGACAAAGCTGAT 1800
Db GCTCAATCTGGCCTGTTGGCTGAATATCAACCGGACAAAGCGTTCCAGGACAAAGCTGAT 1800
1801 CCGTTTACCTGGATGCGGTACGTTTACACCGCAAGCTGATGCTTACCCGATCCCTGTT 1860
Db CCGTTTACCTGGATGCGGTACGTTTACACCGCAAGCTGATGCTTACCCGATCCCTGTT 1860
1861 GAAGGTTTATCGCTGATTTATACCAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920
Db GAAGGTTTATCGCTGATTTATACCAAGATCTGCTGCCGAACCCGCCAAAACCTGGGAA 1920

1921 GAGATCCGCGCTGGATAAAGAACTGAAACCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
Db GAGATCCGCGCTGGATAAAGAACTGAAACCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
1981 CTGCAAGAAACCGTACTTCCACCTGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db CTGCAAGAAACCGTACTTCCACCTGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
2041 TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGCGAAAGCG 2100
Db TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGCGAAAGCG 2100
2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACAACATGTAATGACAGACCGATTAC 2160
Db GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACAACATGTAATGACAGACCGATTAC 2160
2161 TCCATCGCAGAAAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db TCCATCGCAGAAAGCTGCTTTTAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
2221 GCATGGTCCAAACATCGACACCAAGTGAATTAATGTTAAACGGTACTGCGGACCTTC 2280
Db GCATGGTCCAAACATCGACACCAAGTGAATTAATGTTAAACGGTACTGCGGACCTTC 2280
2281 AAGGTCACACCATCAAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db AAGGTCACACCATCAAAACCGTTGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
2341 CCGAACAAAGAGCTGGGAAAAGAGTTCTCCGAAACATATCTGCTGACGTGATGAAGTCTG 2400
Db CCGAACAAAGAGCTGGGAAAAGAGTTCTCCGAAACATATCTGCTGACGTGATGAAGTCTG 2400
2401 GAAAGCGTTTAAATAAGACAAACCGCTGGGTGCGCTAGCGCTCAAGTCTTACGAGGAAG 2460
Db GAAAGCGTTTAAATAAGACAAACCGCTGGGTGCGCTAGCGCTCAAGTCTTACGAGGAAG 2460
2461 TTGGGAAAAGATCCACGATTTGCGCCACCATGGAACCGCCAGAAAGGTGAAATCATG 2520
Db TTGGGAAAAGATCCACGATTTGCGCCACCATGGAACCGCCAGAAAGGTGAAATCATG 2520
2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTGATGCGCTGCTGCTGCTGATCAACGCC 2580
Db CCGAACATCCCGCAGATGTCGCTTTCTGTTGATGCGCTGCTGCTGCTGATCAACGCC 2580
2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
2641 AACAAACAACAATAACAATAACAACACCTCGGGATCGAGGGAAGGATTTCAGAAATTC 2700
Db AACAAACAACAATAACAATAACAACACCTCGGGATCGAGGGAAGGATTTCAGAAATTC 2700
2701 GGATCTCTTCTCTGTTGGCCGAGCGGCTTCGAGCCCGG 2741
Db GGATCTCTTCTCTGTTGGCCGAGCGGCTTCGTTGGGAGG 2741

RESULT 7
AX172306
LOCUS AX172306 8101 bp DNA linear PAT 03-JUL-2001
DEFINITION Sequence 5 from Patent WO014467.
ACCESSION AX172306
VERSION AX172306.1 GI:14597487
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Goubin-Gramatica,F., Ducommun,B. and Prevost,G.
TITLE Method for obtaining human cdc25 phosphatases and method for identifying human cdc25 phosphatase modulators
JOURNAL Patent: WO 014467-A 5 21-JUN-2001;
SOCIETE DE CONSEILS DE RECHERCHES ET D'APPLICATIONS SCIENTIFIQUES

FEATURES		(FR)		Location/Qualifiers	
source		1..8101		/organism="synthetic construct"	
		/mol_type="unassigned DNA"			
		/db_xref="taxon:32630"			
		/note="bouche JM109 / pMAL - Hs Cdc25C"			
ORIGIN					
Query Match 82.0%; Score 2705.8; DB 6; Length 8101;					
Best Local Similarity 99.9%; Pred. No. 0;					
Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;					
QY	1	CGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCCCGGAAGAGA	60		
DB	1	CGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCCCGGAAGAGA	60		
QY	61	GTCAATTGAGGTGGTGAATGTGAACCAAGTAACTTATACGATGTCGAGATATGCCG	120		
DB	61	GTCAATTGAGGTGGTGAATGTGAACCAAGTAACTTATACGATGTCGAGATATGCCG	120		
QY	121	GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACCGAGCCAGCCAGCTTTCTCGGAAA	180		
DB	121	GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACCGAGCCAGCCAGCTTTCTCGGAAA	180		
QY	181	CGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTAATTCCTCCCAACCGCGTGGCAC	240		
DB	181	CGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTAATTCCTCCCAACCGCGTGGCAC	240		
QY	241	AACAACTGGCGGGAACAGTGTGCTGATTTGGCGTTGCCACTCCAGCTTGGCCCTGC	300		
DB	241	AACAACTGGCGGGAACAGTGTGCTGATTTGGCGTTGCCACTCCAGCTTGGCCCTGC	300		
QY	301	ACGCCCGTCGCAATTTGTCGGCGGATTAATCTCGCGCGATCACTGGTGGCCAGCG	360		
DB	301	ACGCCCGTCGCAATTTGTCGGCGGATTAATCTCGCGCGATCACTGGTGGCCAGCG	360		
QY	361	TGTTGGTTCGATGGTAGAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCACAATC	420		
DB	361	TGTTGGTTCGATGGTAGAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCACAATC	420		
QY	421	TTCTCGCGCAACGGTTCAGTGGCTGATCACTAATCTCCCGTGAAGACCGAGATGCCA	480		
DB	421	TTCTCGCGCAACGGTTCAGTGGCTGATCACTAATCTCCCGTGAAGACCGAGATGCCA	480		
QY	481	TTGCTGTGGAAGTGCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCCTGACCA	540		
DB	481	TTGCTGTGGAAGTGCCTGCACATAATGTTCCGGCGTTATTTCTTGATGTCCTGACCA	540		
QY	541	CACCATCAACAGTATTAATTTCTCCCATGAAGACCGTACGCGACTGGGGTGGAGCATC	600		
DB	541	CACCATCAACAGTATTAATTTCTCCCATGAAGACCGTACGCGACTGGGGTGGAGCATC	600		
QY	601	TGGTCGATTCGGTTCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCCTCTCGG	660		
DB	601	TGGTCGATTCGGTTCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCCTCTCGG	660		
QY	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720		
DB	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720		
QY	721	CGGAACGGGAAGGCGATGGAGTGGCATGTCGGTTTTTCAACCAACCAATGCTGA	780		
DB	721	CGGAACGGGAAGGCGATGGAGTGGCATGTCGGTTTTTCAACCAACCAATGCTGA	780		
QY	781	ATGAGGGCATGTTTCCCACTCGCATGCTGGTTGGCAACGATCAGATGGCGTGGCGCAA	840		
DB	781	ATGAGGGCATGTTTCCCACTCGCATGCTGGTTGGCAACGATCAGATGGCGTGGCGCAA	840		
QY	841	TGCGCGCATTCAGGATTCGGGCTGCGGTGGTGGCGATATCTCGGTAGTGGGATACG	900		
DB	841	TGCGCGCATTCAGGATTCGGGCTGCGGTGGTGGCGATATCTCGGTAGTGGGATACG	900		
QY	901	ACGATACCGAAGACAGCTCATGTTATATATCCCGCGTTAAACACCATCAAAACAGGATTTTC	960		
DB	901	ACGATACCGAAGACAGCTCATGTTATATATCCCGCGTTAAACACCATCAAAACAGGATTTTC	960		
QY	961	GCCTGCTGGGGCAAAACAGCGTGCACCGCTTGTCTGCAACTCTCTCAGGGCCAGCGGTGA	1020		
DB	961	GCCTGCTGGGGCAAAACAGCGTGCACCGCTTGTCTGCAACTCTCTCAGGGCCAGCGGTGA	1020		
QY	1021	AGGCAATACGCTGTTGCCGTCTCACTGTGTGAAGAAAAACACCTTGGCGCCCAATA	1080		
DB	1021	AGGCAATACGCTGTTGCCGTCTCACTGTGTGAAGAAAAACACCTTGGCGCCCAATA	1080		
QY	1081	CGCAACCGCTCTCCCGCGGTGGCGGATTCATTAAATGCAGCTGGCAGCAGGTTT	1140		
DB	1081	CGCAACCGCTCTCCCGCGGTGGCGGATTCATTAAATGCAGCTGGCAGCAGGTTT	1140		
QY	1141	CCCGACTGGAAGCGGCGAGTGAGCGCAAGCGCAATTAATGTAGTTAGCTCACTATTAG	1200		
DB	1141	CCCGACTGGAAGCGGCGAGTGAGCGCAAGCGCAATTAATGTAGTTAGCTCACTATTAG	1200		
QY	1201	GCACAATTCATGTTTGAACAGCTTATCATCGACTGCAACCGGTGCAACAAATGCTTGGCG	1260		
DB	1201	GCACAATTCATGTTTGAACAGCTTATCATCGACTGCAACCGGTGCAACAAATGCTTGGCG	1260		
QY	1261	TCAGGAGCCATCGAAAGCTGTGTATGGCTGTGAGCTGTAATCACTGCAATAATTCG	1320		
DB	1261	TCAGGAGCCATCGAAAGCTGTGTATGGCTGTGAGCTGTAATCACTGCAATAATTCG	1320		
QY	1321	TGTCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTCGCGCCGACATCAAAACGGTT	1380		
DB	1321	TGTCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTCGCGCCGACATCAAAACGGTT	1380		
QY	1381	CTGGCAATATCTGAAATGAGCTGTGTGACAAATTAATCATCGGCTCGTATAATGTGTGGA	1440		
DB	1381	CTGGCAATATCTGAAATGAGCTGTGTGACAAATTAATCATCGGCTCGTATAATGTGTGGA	1440		
QY	1441	ATTGTGAGCGGATAACAAATTTCAACAGGAAACAGCCAGTCCGTTTATAGGTGTTTTACGA	1500		
DB	1441	ATTGTGAGCGGATAACAAATTTCAACAGGAAACAGCCAGTCCGTTTATAGGTGTTTTACGA	1500		
QY	1501	GCACTTCACAAACAGGACCATAGATTATGAAATCGAAGAGGTAACTGGTAAATCTGG	1560		
DB	1501	GCACTTCACAAACAGGACCATAGATTATGAAATCGAAGAGGTAACTGGTAAATCTGG	1560		
QY	1561	ATTAAACGGCGATAAAGCTATACCGTCTCGCTGAGTTCGGTAAGAAATTCGAGAAAGAT	1620		
DB	1561	ATTAAACGGCGATAAAGCTATACCGTCTCGCTGAGTTCGGTAAGAAATTCGAGAAAGAT	1620		
QY	1621	ACCGGAATTAAGTTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCACAGTT	1680		
DB	1621	ACCGGAATTAAGTTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCACAGTT	1680		
QY	1681	CGGGCAACTGGCGATGGCCCTGACATTATCTTGGGCAACGACCGCTTTGGTGGCTAC	1740		
DB	1681	CGGGCAACTGGCGATGGCCCTGACATTATCTTGGGCAACGACCGCTTTGGTGGCTAC	1740		
QY	1741	GCTCAATCTGGCTGTGGCTGAAATCACCCCGCAAAAGCGTTCCAGGACAGCTGTAT	1800		
DB	1741	GCTCAATCTGGCTGTGGCTGAAATCACCCCGCAAAAGCGTTCCAGGACAGCTGTAT	1800		
QY	1801	CCGTTTACCTGGGATCCGTTACAAAGATCTGATTCCTTACCCGATCGCTGTT	1860		
DB	1801	CCGTTTACCTGGGATCCGTTACAAAGATCTGATTCCTTACCCGATCGCTGTT	1860		
QY	1861	GAGCGTTATCGCTGATTTATAAAGATCTGCTCCGAAACCCGCGCAAAACCTGGGAA	1920		
DB	1861	GAGCGTTATCGCTGATTTATAAAGATCTGCTCCGAAACCCGCGCAAAACCTGGGAA	1920		
QY	1921	GAGATCCCGCGCTGATTAAGAACTGAAAGCGGAAAGGTAAAGCGCGCTGTATGTTCAAC	1980		
DB	1921	GAGATCCCGCGCTGATTAAGAACTGAAAGCGGAAAGGTAAAGCGCGCTGTATGTTCAAC	1980		
QY	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATTTGCTGTGACGGGGGTTATGCGTTCAAG	2040		

Db	1981	CTGCAGAAACCGTACTTTCACCTGGCCGCTGATTGCTGACGGGGTTATCGTTCAAG	2040
Qy	2041	TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGGAAAGCG	2100
Db	2041	TATGAAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATTAACGCTGGCGGAAAGCG	2100
Qy	2101	GGTCTGACCTTCTCGTGGTGAACCTGATTAATAAACAACAACACATGAATGCGAGACACCGATTAC	2160
Db	2101	GGTCTGACCTTCTCGTGGTGAACCTGATTAATAAACAACAACACATGAATGCGAGACACCGATTAC	2160
Qy	2161	TCCATCGCAGAAAGCTGCTTTTAAAGCGCAACAGCGATGACCATCAAGCGCCCGTGG	2220
Db	2161	TCCATCGCAGAAAGCTGCTTTTAAAGCGCAACAGCGATGACCATCAAGCGCCCGTGG	2220
Qy	2221	GCATGGTCCAAATCGACACACGCAAGTGAATTAATGGTGAACGGTACTGCGGACCTTC	2280
Db	2221	GCATGGTCCAAATCGACACACGCAAGTGAATTAATGGTGAACGGTACTGCGGACCTTC	2280
Qy	2281	AAGGTCAACCATCAACACCGTTCGTTGGCGTGTGAGGCGAGGTATTAAACGCCCAAGT	2340
Db	2281	AAGGTCAACCATCAACACCGTTCGTTGGCGTGTGAGGCGAGGTATTAAACGCCCAAGT	2340
Qy	2341	CGAAACAAAGAGCTGCGCAAAAGAGTTCTTCGAAACTATCTGCTGACTGATGAAGTCTG	2400
Db	2341	CGAAACAAAGAGCTGCGCAAAAGAGTTCTTCGAAACTATCTGCTGACTGATGAAGTCTG	2400
Qy	2401	GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG	2460
Db	2401	GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG	2460
Qy	2461	TTGGCGAAAGATCCACGATTAATGCGCGCACCATGGAAGAACGCCAGAAAGGTGAATCATG	2520
Db	2461	TTGGCGAAAGATCCACGATTAATGCGCGCACCATGGAAGAACGCCAGAAAGGTGAATCATG	2520
Qy	2521	CGAACATCCCGCAGATGTCGCTTCTGTTGATGCGGTGCTGCTGATCAACGCC	2580
Db	2521	CGAACATCCCGCAGATGTCGCTTCTGTTGATGCGGTGCTGCTGATCAACGCC	2580
Qy	2581	GCCAGCGGTGTCAGACATGTCGATGAAGCCCTGAAAGAGCGCGACACTAAATTCGAGCTCG	2640
Db	2581	GCCAGCGGTGTCAGACATGTCGATGAAGCCCTGAAAGAGCGCGACACTAAATTCGAGCTCG	2640
Qy	2641	AACAAACAACAATAACAATAACAACACCTCGGGATCGAGGAAAGGATTCAGAAATTC	2700
Db	2641	AACAAACAACAATAACAATAACAACACCTCGGGATCGAGGAAAGGATTCAGAAATTC	2700
Qy	2701	GGATCCTCT	2709
Db	2701	GGATCCTCT	2709

RESULT 8
AX377531

LOCUS AX377531 6648 bp DNA linear PAT 18-MAR-2002

DEFINITION Sequence 8 from Patent W00212553.

ACCESSION AX377531

VERSION AX377531.1 GI:19573717

KEYWORDS

SOURCE synthetic construct

ORGANISM synthetic construct

REFERENCE 1 other sequences; artificial sequences.

AUTHORS Kappel, A., Polakowski, T., Pignot, M., Windhab, N., Behrens, H., and Muth, J.

TITLE Method for detecting mutations in nucleotide sequences

JOURNAL Patent: WO 0212553-A 8 14-FEB-2002;

FEATURES

Location/Qualifiers

1. .6648

/organism="synthetic construct"

/mol_type="unassigned DNA"

/db_xref="taxon:32630"

ORIGIN	/note="Plasmid pMAL-c2x"			
Query Match	81.8%	Score 2701;	DB 6;	Length 6648;
Best Local Similarity	99.8%	Pred. No. 0;		
Matches 2704;	Conservative	0;	Mismatches	5; Indels 0; Gaps 0;
Qy	1	CCGACACCATCGAATGGTGTGCAAAACCTTTCCGGGTATGGCATGATAGGCCCGGGAAGAGA	60	
Db	1	CCGACACCATCGAATGGTGTGCAAAACCTTTCCGGGTATGGCATGATAGGCCCGGGAAGAGA	60	
Qy	61	GTCAATTCCAGGTGTGAATGTGAACCCAGTAAAGCTTATACGATGTCGACAGATATCGG	120	
Db	61	GTCAATTCCAGGTGTGAATGTGAACCCAGTAAAGCTTATACGATGTCGACAGATATCGG	120	
Qy	121	GTGTCTCTTATCAGACCGTTCCTCCGGGTGTGTAACACGAGCCAGCCACGTTTCTGCGAAAA	180	
Db	121	GTGTCTCTTATCAGACCGTTCCTCCGGGTGTGTAACACGAGCCAGCCACGTTTCTGCGAAAA	180	
Qy	181	CGCGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTACATTCGCAACCGCTGGCAC	240	
Db	181	CGCGGAAAAAGTGAAGCGCGGATGGCGGAGCTGAATTACATTCGCAACCGCTGGCAC	240	
Qy	241	AACAACTGGCGGCAAAACAGTCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300	
Db	241	AACAACTGGCGGCAAAACAGTCTGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300	
Qy	301	ACGCGCGCTCGCAAAATTTGTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCGCAGCG	360	
Db	301	ACGCGCGCTCGCAAAATTTGTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGCGCAGCG	360	
Qy	361	TGGTGGTTCGATGGTGAAGACGAGCGCGCTGGAAGCCTGTAAAGCGGCGGTGCAATTC	420	
Db	361	TGGTGGTTCGATGGTGAAGACGAGCGCGCTGGAAGCCTGTAAAGCGGCGGTGCAATTC	420	
Qy	421	TTCTCGCGCAACGCTGAGTGGGCTGATTAATTAATCTCCGTCGATGATCAACAGGATGCCA	480	
Db	421	TTCTCGCGCAACGCTGAGTGGGCTGATTAATTAATCTCCGTCGATGATCAACAGGATGCCA	480	
Qy	481	TTGCTGTGGAAGCTGCTGCACTAAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA	540	
Db	481	TTGCTGTGGAAGCTGCTGCACTAAATGTTCCGGCGTTATTTCTTGATGCTCTGACCCAGA	540	
Qy	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGCTACGCGACTGGCGGTGGAGCATC	600	
Db	541	CACCCATCAACAGTATTAATTTCTCCCATGAAGACGCTACGCGACTGGCGGTGGAGCATC	600	
Qy	601	TGGTGGCATGGTCCAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG	660	
Db	601	TGGTGGCATGGTCCAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG	660	
Qy	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720	
Db	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720	
Qy	721	CGGAAACGGAAGCGGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA	780	
Db	721	CGGAAACGGAAGCGGACTGGAGTGCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA	780	
Qy	781	ATGAGGGCATGTTCCCACTGGATGCTGTTGTCGCAACGATCAGATGGCGTGGCGGCA	840	
Db	781	ATGAGGGCATGTTCCCACTGGATGCTGTTGTCGCAACGATCAGATGGCGTGGCGGCA	840	
Qy	841	TGCGGCGCATTAACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATAG	900	
Db	841	TGCGGCGCATTAACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATAG	900	
Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTGTAAACCAATCAAAACAGGATTTTC	960	
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTGTAAACCAATCAAAACAGGATTTTC	960	
Qy	961	GCCTGCTGGGCAACACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGCCAGGCGGTGA	1020	
Db	961	GCCTGCTGGGCAACACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGCCAGGCGGTGA	1020	

```
Db      961  GCCTGCTGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
Qy      1021  AGGGCAATCAGCTGTGGCCCGTCTCACTGCTGAAAGAAACCAACCTGGCGCCCAATA 1080
Db      1021  AGGGCAATCAGCTGTGGCCCGTCTCACTGCTGAAAGAAACCAACCTGGCGCCCAATA 1080
Qy      1081  CGGAACCGGCTCTCCCGCGGCTGGCCGATTCATTAATGAGCTGGCAGCAGAGTTT 1140
Db      1081  CGGAACCGGCTCTCCCGCGGCTGGCCGATTCATTAATGAGCTGGCAGCAGAGTTT 1140
Qy      1141  CCGACTGGAAGCGGAGTGGCGCAACGCAATTAATGAGCTAGCTACTCATTTAG 1200
Db      1141  CCGACTGGAAGCGGAGTGGCGCAACGCAATTAATGAGCTAGCTACTCATTTAG 1200
Qy      1201  GCACAATCTCATTTGACAGCTTATCATGCACTGACGCTGCAACCAATGCTTCGCG 1260
Db      1201  GCACAATCTCATTTGACAGCTTATCATGCACTGACGCTGCAACCAATGCTTCGCG 1260
Qy      1261  TCAGGAGCATCGGAAGCTGGTATGCTGTGAGCTCGTAATCATCTGCAATTCG 1320
Db      1261  TCAGGAGCATCGGAAGCTGGTATGCTGTGAGCTCGTAATCATCTGCAATTCG 1320
Qy      1321  TGTGCTCAAGCGCACTCCCGTCTGATATGTTTTCGCGCAATCATACAGGTT 1380
Db      1321  TGTGCTCAAGCGCACTCCCGTCTGATATGTTTTCGCGCAATCATACAGGTT 1380
Qy      1381  CTGGCAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATTCG 1440
Db      1381  CTGGCAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATTCG 1440
Qy      1441  ATTGTGAGGGATACAAATTTACACAGGAACAGCCAGTCGTTTAGGTTTTCAGGA 1500
Db      1441  ATTGTGAGGGATACAAATTTACACAGGAACAGCCAGTCGTTTAGGTTTTCAGGA 1500
Qy      1501  GCACCTTCCAAACAGGACCATAGATTATGAAACTGGAAGAGTAAACTGCTAATCTGG 1560
Db      1501  GCACCTTCCAAACAGGACCATAGATTATGAAACTGGAAGAGTAAACTGCTAATCTGG 1560
Qy      1561  ATTAACGGGATTAAGGCTATAACCGGCTCGCTGAACTCGGTGAGAAATTCGAGAAAGAT 1620
Db      1561  ATTAACGGGATTAAGGCTATAACCGGCTCGCTGAACTCGGTGAGAAATTCGAGAAAGAT 1620
Qy      1621  ACCGGAATTAAGTCACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCGACAGGTT 1680
Db      1621  ACCGGAATTAAGTCACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTCGACAGGTT 1680
Qy      1681  GCGCAACTGCGGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db      1681  GCGCAACTGCGGATGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy      1741  GCTCAATCTGGCCCTGTTGGCTGAAATCAACCGGACAAAGGTTCCAGGCAAGCTGTAT 1800
Db      1741  GCTCAATCTGGCCCTGTTGGCTGAAATCAACCGGACAAAGGTTCCAGGCAAGCTGTAT 1800
Qy      1801  CCGTTTACCTGGGATGCGGTACGTTACAAACGGCAAGCTGATTGCTTACCGATCGCTGT 1860
Db      1801  CCGTTTACCTGGGATGCGGTACGTTACAAACGGCAAGCTGATTGCTTACCGATCGCTGT 1860
Qy      1861  GAAGGTTATCGCTGATTATAACAAAGATCTGCTGCGCAACCGCCCAAAACCTGGGAA 1920
Db      1861  GAAGGTTATCGCTGATTATAACAAAGATCTGCTGCGCAACCGCCCAAAACCTGGGAA 1920
Qy      1921  GAGATCCGGCGCTGGATTAAGAACTGAAAGCGAAAGTAAAGCGGCTGATGTTTCAAC 1980
Db      1921  GAGATCCGGCGCTGGATTAAGAACTGAAAGCGAAAGTAAAGCGGCTGATGTTTCAAC 1980
Qy      1981  CTGCAAGAACCGTACTTCACTGGCCGCTGATTCGCTGACGGGGGTATGCGTTCAAG 2040
Db      1981  CTGCAAGAACCGTACTTCACTGGCCGCTGATTCGCTGACGGGGGTATGCGTTCAAG 2040
Qy      2041  TATGAAACCGCAAGTACGACATTAAGACGTTGGGCGTGGATACGCTGGCGGAAAGCG 2100
Db      2041  TATGAAACCGCAAGTACGACATTAAGACGTTGGGCGTGGATACGCTGGCGGAAAGCG 2100
```

```
Qy      2101  GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACAACATGAATGCAGACACCGATTAC 2160
Db      2101  GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACAACATGAATGCAGACACCGATTAC 2160
Qy      2161  TCCATCGCAGAGCTGCTTTTAAAGAGCGAAACAGCATGACCATCAACGCGCCCGTGG 2220
Db      2161  TCCATCGCAGAGCTGCTTTTAAAGAGCGAAACAGCATGACCATCAACGCGCCCGTGG 2220
Qy      2221  GCATGGTCCAAACATCGACACAGCAAAAGTGAATTTATGGTGTAAACGTAATTCGCGACCTTC 2280
Db      2221  GCATGGTCCAAACATCGACACAGCAAAAGTGAATTTATGGTGTAAACGTAATTCGCGACCTTC 2280
Qy      2281  AAGGCTCAACCATCAACCGCTTGGTGGCGTGTGAGCGCAGGTATTAACGCGCCAGT 2340
Db      2281  AAGGCTCAACCATCAACCGCTTGGTGGCGTGTGAGCGCAGGTATTAACGCGCCAGT 2340
Qy      2341  CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTTCTG 2400
Db      2341  CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTTCTG 2400
Qy      2401  GAAGCGTTTAAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Db      2401  GAAGCGTTTAAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Qy      2461  TTGCGGAAAGATCCACGCTATTGCGCCACCATGGAACCGCCAGAAAGGTTGAATCATG 2520
Db      2461  TTGCGGAAAGATCCACGCTATTGCGCCACCATGGAACCGCCAGAAAGGTTGAATCATG 2520
Qy      2521  CCGAACATCCGCGAGATGTCGCTTTCTGGTATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Db      2521  CCGAACATCCGCGAGATGTCGCTTTCTGGTATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2580
Qy      2581  GCCAGCGCTGTCAGACTGTCGATGAAGCCCTGAAAGACCGCAGACGACTAATTCGAGCTCG 2640
Db      2581  GCCAGCGCTGTCAGACTGTCGATGAAGCCCTGAAAGACCGCAGACGACTAATTCGAGCTCG 2640
Qy      2641  AACAAACAAACAAATAACAAATAACAAATAACAAATAACAAATAACAAATAACAAATAAC 2700
Db      2641  AACAAACAAACAAATAACAAATAACAAATAACAAATAACAAATAACAAATAACAAATAAC 2700
Qy      2701  GGATCTCTCT 2709
Db      2701  GGATCTCTCT 2709

RESULT 9
LOCUS      A91965      7475 bp      DNA      circular      PAT 22-JAN-2000
DEFINITION Sequence 1 from Patent WO9822496.
ACCESSION  A91965
VERSION     A91965.1  GI:6740811
KEYWORDS   'unidentified
SOURCE      unidentified
ORGANISM    unidentified
REFERENCE   1 (bases 1 to 7475)
AUTHORS     Attwood,M.R. and Hurst,D.N.
TITLE       ANTIVIRAL PEPTIDE DERIVATIVES
JOURNAL     Patent: WO 9822496-A 1 28-MAY-1998;
            HOFFMANN LA ROCHE (CH)
FEATURES    Location/Qualifiers
            source          1..7475
                        /organism="unidentified"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:32644"
ORIGIN
Query Match      81.8%; Score 2700; DB 6; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1  CCGACACCATCGAATCGTGTCAAAACCTTTTCGCGGTATGCGATGATAGCGCCCGGAAAGAGA 60
```

Db 1 CCGACACCATGAAATGGTGCAGAAACCTTTTCGCGGTATGGCATATGATAGCGCCCGGAAGAGA 60
Qy 61 GTCAATTCAGGGTGTGAATGTGAACACAGTAAACGTTATACGATGTCCGAGAGTATGCCG 120
Db 61 GTCAATTCAGGGTGTGAATGTGAACACAGTAAACGTTATACGATGTCCGAGAGTATGCCG 120
Qy 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCACGTTTCTCGGAAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACACGAGCCAGCCACGTTTCTCGGAAAA 180
Qy 181 CGCGGAAAAAGTGAACGCGCATGGCGAGCTGAATTAACATCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAACGCGCATGGCGAGCTGAATTAACATCCCAACCGCGTGGCAC 240
Qy 241 AACAACTGCGCGGCAAAACAGTGTGCTGATTTGGCGTTGCGCAACCTCCAGTCTGGGCCCTGC 300
Db 241 AACAACTGCGCGGCAAAACAGTGTGCTGATTTGGCGTTGCGCAACCTCCAGTCTGGGCCCTGC 300
Qy 301 ACGGCGCGTCGCAAAATGTGCGGCGCATTAATCTCGCGCGCATCAACTGGGTGCGCAGCG 360
Db 301 ACGGCGCGTCGCAAAATGTGCGGCGCATTAATCTCGCGCGCATCAACTGGGTGCGCAGCG 360
Qy 361 TGGTGTGTCGATGCTAGTGAACGAGCGCGTCGAAGCCTGTAAAGCGCGGTGCAACAATC 420
Db 361 TGGTGTGTCGATGCTAGTGAACGAGCGCGTCGAAGCCTGTAAAGCGCGGTGCAACAATC 420
Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTATCCGTCGATGATCACCAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTATCCGTCGATGATCACCAGGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCCTGCACTAAATGTTCGGCGGTATTTCTTGATGTCTGACACAGA 540
Db 481 TTGCTGTGGAAGCTGCCTGCACTAAATGTTCGGCGGTATTTCTTGATGTCTGACACAGA 540
Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGCATGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGCATGGCGGTGGAGCATC 600
Qy 601 TGGTGCATTTGGTCAACAGCAAAATCGCGCTGTAGCGGCGCCATTAAGTTCTGTCTCGG 660
Db 601 TGGTGCATTTGGTCAACAGCAAAATCGCGCTGTAGCGGCGCCATTAAGTTCTGTCTCGG 660
Qy 661 CGCGTCTGCTGCTGGCTGGCATAAATACTCACTCGCAATCAAAATTCAGCGCATAG 720
Db 661 CGCGTCTGCTGCTGGCTGGCATAAATACTCACTCGCAATCAAAATTCAGCGCATAG 720
Qy 721 CGGAACGGGAAGGCACTGGAGTGCATGTCCGGTTTCAACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGCACTGGAGTGCATGTCCGGTTTCAACAAACCATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCGCATGGCGCTGGGCGCAA 840
Db 781 ATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCGATGGCGCTGGGCGCAA 840
Qy 841 TGCGCGCCATTACCGAGTCCGGGCTGGCGTTTGGTGGGATATCTCGGTAGTGGGATAG 900
Db 841 TGCGCGCCATTACCGAGTCCGGGCTGGCGTTTGGTGGGATATCTCGGTAGTGGGATAG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAATCAAAAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAAACCAATCAAAAGGATTTTC 960
Qy 961 GCCTGCTGGGCAACACGAGCGTGGACCGCTTCTGCACTCTCTCAGGCGCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGCGTGGACCGCTTCTGCACTCTCTCAGGCGCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAAACCACTGGCGGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGGCCGCTCTCACTGGTGAAGAAAAACCACTGGCGGCCCAATA 1080
Qy 1081 CGCAACCGCCTCTCCCGCGGTTGGCGCATTCATTAATGAGCTGGCACACAGGTTT 1140

Db 1081 CGCAACCGCCTCTCCCGCGGTTGGCGATTCATTAATGACAGTGGCACGACAGGTTT 1140
Qy 1141 CCCGACTGGAAAGCGGCGAGTGAAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAAGCGGCGAGTGAAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Qy 1201 GCACAAATCTCATGTTTGAAGCTTATCATGCACTGACGCGTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGAAGCTTATCATGCACTGACGCGTGACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGACAGCCATCGGAAGCTGTGATGCGTGTGAGTGTGAGTGTGATTAATCACTGCATAATTCG 1320
Db 1261 TCAGGACAGCCATCGGAAGCTGTGATGCGTGTGAGTGTGAGTGTGATTAATCACTGCATAATTCG 1320
Qy 1321 TGTGCTCAAGCGCGACCTCCCGTTCTGGAATATGTTTTCGCGCGACATCATAAACGGTT 1380
Db 1321 TGTGCTCAAGCGCGACCTCCCGTTCTGGAATATGTTTTCGCGCGACATCATAAACGGTT 1380
Qy 1381 CTGGCAAAATTTCTGAAATGAGCTTGCACAAATTAATCATCGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTTGCACAAATTAATCATCGCTCGTATATGTTGGA 1440
Qy 1441 ATTGTGAGCGGATAACAAATTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTTCAACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
Qy 1501 GCACCTTCAACAAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAATCTCG 1560
Db 1501 GCACCTTCAACAAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAATCTCG 1560
Qy 1561 ATTTAACGGCGATAAAGGCTATAACGGTCTCGCTGAGTTCGGTAAAGAAATTCGAGAAAGT 1620
Db 1561 ATTTAACGGCGATAAAGGCTATAACGGTCTCGCTGAGTTCGGTAAAGAAATTCGAGAAAGT 1620
Qy 1621 ACCGGAATTTAAAGTCACCGTTGAGCATCCGATATAACTGGAAGAGAAATTCGCCAGGTT 1680
Db 1621 ACCGGAATTTAAAGTCACCGTTGAGCATCCGATATAACTGGAAGAGAAATTCGCCAGGTT 1680
Qy 1681 CGGCAACTGCGCATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGCGCATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCGGACAAAGCGTTCACGACAGCTGTAT 1800
Db 1741 GCTCAATCTGCGCTGTTGGCTGAAATCAACCGGACAAAGCGTTCACGACAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTTGCTTACCCGATCGCTGT 1860
Qy 1861 GAAGCGTTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAACCCGCCAAAAAACCCTGGAA 1920
Db 1861 GAAGCGTTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAACCCGCCAAAAAACCCTGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCCACCTGGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCCACCTGGCGCTGATTTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAAACGCAAGTACGACATTAAGAACGCTGGCGGTGGATTAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGAACGCTGGCGGTGGATTAACGCTGGCGCGAAAGCG 2100
Qy 2101 GGTCTGAACCTTCTGCTGTTGAACCTGATTAATAACAAACATGAATGACAGACACCGATTAC 2160
Db 2101 GGTCTGAACCTTCTGCTGTTGAACCTGATTAATAACAAACATGAATGACAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCGCTTTTAAATGAAGCGAAACAGCGATGCCATCAACGCGCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCGCTTTTAAATGAAGCGAAACAGCGATGCCATCAACGCGCGCTGG 2220

QY 1321 TGTGCTCAAGGGGCACTCCCGTTCTGGATAAATGTTTTTGGCCGCAATCATATAACGGTT 1380
Db 1321 TGTGCTCAAGGGGCACTCCCGTTCTGGATAAATGTTTTTGGCCGCAATCATATAACGGTT 1380
QY 1381 CTGGCAAAATATTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAAAATATTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
QY 1441 ATTGTAGCGGATAACAAATTTACACAGGAAACAGCCAGTCGCTTTAGGTGTTTTTCACGA 1500
Db 1441 ATTGTAGCGGATAACAAATTTACACAGGAAACAGCCAGTCGCTTTAGGTGTTTTTCACGA 1500
QY 1501 GCACCTCACCAACAGGACCATAGATTATGAAAACTGAAGAAGGTAAACTGGTAATCTGG 1560
Db 1501 GCACCTCACCAACAGGACCATAGATTATGAAAACTGAAGAAGGTAAACTGGTAATCTGG 1560
QY 1561 ATTTAAGCGGATAAAGGCTATAACGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTTAAGCGGATAAAGGCTATAACGCTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCACAGGTT 1680
QY 1681 GGGCAACTGGCGATGGCCCTGACATTATCTCTGGGCAACGACCGCTTTGGTGGCTAC 1740
Db 1681 GGGCAACTGGCGATGGCCCTGACATTATCTCTGGGCAACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGTGCTGAAATCACCCGGACAAAGCGTTCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGTGCTGAAATCACCCGGACAAAGCGTTCAGGACAAAGCTGTAT 1800
QY 1801 CGGTTTACCTGGGATGCGGTACGTTACAAAGCAAGCTGATGCTTTACCGCATCGCTGT 1860
Db 1801 CGGTTTACCTGGGATGCGGTACGTTACAAAGCAAGCTGATGCTTTACCGCATCGCTGT 1860
QY 1861 GAAGCGTTATCGTGATTTATACAAAGATCTGCTGCGAAACCGGCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGTGATTTATACAAAGATCTGCTGCGAAACCGGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCTGGATAAGCAATGAAGCGAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCTGGATAAGCAATGAAGCGAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACCGCAAGTACACATTAAGACGTTGGGCGTGGATAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAAACCGCAAGTACACATTAAGACGTTGGGCGTGGATAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACACATGATGACACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACACATGATGACACACCGATTAC 2160
QY 2161 TCCATCGCAAGAGCTGCTTTAATAAGCGGAAACAGCGCATGACCATCAACGCGCCGCTG 2220
Db 2161 TCCATCGCAAGAGCTGCTTTAATAAGCGGAAACAGCGCATGACCATCAACGCGCCGCTG 2220
QY 2221 GCATGCTCAACATCGACACCGAAGTGAATTTATGGTGAACGCGTACTGCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACCGAAGTGAATTTATGGTGAACGCGTACTGCGGACCTTC 2280
QY 2281 AAGGTCACCATCCAAACCGTTGCTGGCTGCTGAGCGCAGGTTATTAACGCGCCAGT 2340
Db 2281 AAGGTCACCATCCAAACCGTTGCTGGCTGCTGAGCGCAGGTTATTAACGCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400

QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTAGCGCTGAACTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTAGCGCTGAACTCTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCACGCTATTGCGCCACCATGGAAAAACGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGCTATTGCGCCACCATGGAAAAACGCCAGAAAGGTGAAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCCGTGCTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCCGTGCTGCTGCTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGCACACTGTCGATGAAGCCCTGAAAGACGCCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGCACACTGTCGATGAAGCCCTGAAAGACGCCAGACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAAACAATAAACAACCTCGGGATCGAGGGAAAGATTTCAGAAATTC 2700
Db 2641 AACAAACAACAATAAACAATAAACAACCTCGGGATCGAGGGAAAGATTTCAGAAATTC 2700

RESULT 11
AR207294
LOCUS AR207294 7475 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 1 from patent US 6372883.
ACCESSION AR207294
VERSION AR207294.1 GI:21506162
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 7475)
AUTHORS Atwood,M,Richard., Hurst,D,Nigel., Jones,P,Stephen.,
Kay,P,Brittain., Raynham,T,Michael. and Wilson,F,Xavier.
TITLE Antiviral medicaments
JOURNAL Patent: US 6372883-A 1 16-APR-2002;
FEATURES
source Location/Qualifiers
1..7475
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Query Match 81.8%; Score 2700; DB 6; Length 7475;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCACCATCGAATGGTGCBAACCTTTCCGGGTATGCGATAGCGCCCGGAAGAGA 60
Db 1 CGCACCATCGAATGGTGCBAACCTTTCCGGGTATGCGATAGCGCCCGGAAGAGA 60
QY 61 GTCAAATTCAGGTTGTTGAATGTGAACCCAGTAACTTATACGATGTCGAGAGTATGCCG 120
Db 61 GTCAAATTCAGGTTGTTGAATGTGAACCCAGTAACTTATACGATGTCGAGAGTATGCCG 120
QY 121 GTGTCCTTTATCAGACCGTTTCCCGCGTGTGAACCCAGCCAGCCAGCTTTCTGGAAAAA 180
Db 121 GTGTCCTTTATCAGACCGTTTCCCGCGTGTGAACCCAGCCAGCCAGCTTTCTGGAAAAA 180
QY 181 CGCGGAAAAAGTGAAGCGGATGCGGAGCTCAATTAATCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGGATGCGGAGCTCAATTAATCCCAACCGCGTGGCAC 240
QY 241 AACAACTGCGGGCAACAGTCTGCTGATTGGGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGCGGGCAACAGTCTGCTGATTGGGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 AGCGCCGTTCGCAAAATTTGTCGGCGGATTAATCTCGGCCCGATCAACTGGGTGCCAGG 360
Db 301 AGCGCCGTTCGCAAAATTTGTCGGCGGATTAATCTCGGCCCGATCAACTGGGTGCCAGG 360
QY 361 TGGTGGTCTGATGCTAGAACGAGCGGCTCGAAGCCTGTAAAGCGGCGTGCACAAATC 420
Db 361 TGGTGGTCTGATGCTAGAACGAGCGGCTCGAAGCCTGTAAAGCGGCGTGCACAAATC 420

QY	421	TTCTCGGCAACCGGTGAGTGGGCTGATCATTAACATATCGCTGGATGACCGAGTGCCA	480
Db	421	TTCTCGGCAACCGGTGAGTGGGCTGATCATTAACATATCGCTGGATGACCGAGTGCCA	480
QY	481	TTGCTGTGGAAGCTGCTGCACTAAATGTTCCCGCGTTATTTCTTGATGCTCTGACCA	540
Db	481	TTGCTGTGGAAGCTGCTGCACTAAATGTTCCCGCGTTATTTCTTGATGCTCTGACCA	540
QY	541	CACCCATCAACAGTATATTTCTCCCATGAAGACGGTACGGATCGGCGTGGAGCATC	600
Db	541	CACCCATCAACAGTATATTTCTCCCATGAAGACGGTACGGAGCTGGGCGTGGAGCATC	600
QY	601	TGCTCGCATTTGGTTCACCGCAAAATCGCGCTGTTAGCGGCGCCATTAAAGTCTGCTCGG	660
Db	601	TGCTCGCATTTGGTTCACCGCAAAATCGCGCTGTTAGCGGCGCCATTAAAGTCTGCTCGG	660
QY	661	CGGCTCTGCGTCTGGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Db	661	CGGCTCTGCGTCTGGCTGGCTGGCAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAACGGGAAGCGACTGAGTGGCCATCTCGGTTTCAACAAACCATGCAATGCTGA	780
Db	721	CGGAACGGGAAGCGACTGAGTGGCCATCTCGGTTTCAACAAACCATGCAATGCTGA	780
QY	781	ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTGCCAAACGATCAGATGGCGCTGGGCGCA	840
Db	781	ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTGCCAAACGATCAGATGGCGCTGGGCGCA	840
QY	841	TGCGGCGCAATACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGGCGCAATACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACACCATCAACAGGATTTTC	960
Db	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACACCATCAACAGGATTTTC	960
QY	961	GCCTGCTGGGGCAACACAGCTGGAACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA	1020
Db	961	GCCTGCTGGGGCAACACAGCTGGAACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA	1020
QY	1021	AGGGCAATCAGCTGTTGCCGCTCACTGTTGAAAGAAACACCCCTGGCGGCCAATA	1080
Db	1021	AGGGCAATCAGCTGTTGCCGCTCACTGTTGAAAGAAACACCCCTGGCGGCCAATA	1080
QY	1081	CGCAACCGCTCTCCCGCGGTTGGCGATTCAATTAATGAGCTGGCAGACAGGTTT	1140
Db	1081	CGCAACCGCTCTCCCGCGGTTGGCGATTCAATTAATGAGCTGGCAGACAGGTTT	1140
QY	1141	CCGACTGGAAGCGGGCAGTGAGCGAAACGCAATTAATGAGTACCTCATTTAG	1200
Db	1141	CCGACTGGAAGCGGGCAGTGAGCGAAACGCAATTAATGAGTACCTCATTTAG	1200
QY	1201	GCACAAATCTCATGTTTGACAGCTTATCATGCTGACGCTGACCAATGCTTCTGGCG	1260
Db	1201	GCACAAATCTCATGTTTGACAGCTTATCATGCTGACGCTGACCAATGCTTCTGGCG	1260
QY	1261	TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTAAATCACTGCATAATTCG	1320
Db	1261	TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCTAAATCACTGCATAATTCG	1320
QY	1321	TGTCGCTCAAGCGCACTCCCGTTCTGGATAATGTTTTCGCGCGACATCAATACGGTT	1380
Db	1321	TGTCGCTCAAGCGCACTCCCGTTCTGGATAATGTTTTCGCGCGACATCAATACGGTT	1380
QY	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA	1440
Db	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA	1440
QY	1441	ATTGTGAGCGGATCAACAAATTTACAGGAAACAGCGAGTCCGTTTAGGTTTTCACGA	1500
Db	1441	ATTGTGAGCGGATCAACAAATTTACAGGAAACAGCGAGTCCGTTTAGGTTTTCACGA	1500

QY	1501	GCATTTACCAACAGGACCATGATTATGAAAACTGAAGAAGGTAAACTGTTAATCTGG	1560
Db	1501	GCATTTACCAACAGGACCATGATTATGAAAACTGAAGAAGGTAAACTGTTAATCTGG	1560
QY	1561	ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCCGATATAACTGGAAGAGAAATTCGCCACAGTT	1680
Db	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCCGATATAACTGGAAGAGAAATTCGCCACAGTT	1680
QY	1681	CGGCAACTGCGGATGCGCTGACATTAATCTTTCTGGGGAACAGACCGCTTTGGTGGCTAC	1740
Db	1681	CGGCAACTGCGGATGCGCTGACATTAATCTTTCTGGGGAACAGACCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGCGCTGTTGGCTGAAATCACCCCGACAAAGGTTCCAGGACAAAGCTGAT	1800
Db	1741	GCTCAATCTGCGCTGTTGGCTGAAATCACCCCGACAAAGGTTCCAGGACAAAGCTGAT	1800
QY	1801	CCGTTTACCTGGGATGCGCTACGTTTACACGGCAAGCTGATTGCTTACCCTGCTGCTGTT	1860
Db	1801	CCGTTTACCTGGGATGCGCTACGTTTACACGGCAAGCTGATTGCTTACCCTGCTGCTGTT	1860
QY	1861	GAGCGTTATCGCTGATTATTAACAAAGATCTGTCGCGAACCCGCCCAAAACCTTGGGAA	1920
Db	1861	GAGCGTTATCGCTGATTATTAACAAAGATCTGTCGCGAACCCGCCCAAAACCTTGGGAA	1920
QY	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATTGTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATTGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGACCGGGGTTATGCGTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGACCGGGGTTATGCGTTCAAG	2040
QY	2041	TATGAAACCGCAAGTACGACATTAAGACGTCGGCGTGATTAACGCTGGCGGAAGCG	2100
Db	2041	TATGAAACCGCAAGTACGACATTAAGACGTCGGCGTGATTAACGCTGGCGGAAGCG	2100
QY	2101	GGTCTGACCTTCTGCTGTTGACCTGATTAAACAAACACATGATGCAGACACCGATTAC	2160
Db	2101	GGTCTGACCTTCTGCTGTTGACCTGATTAAACAAACACATGATGCAGACACCGATTAC	2160
QY	2161	TCCATGCGAGAACCTGCTTTTAAAGCGAAACAGCGATGACCATCAACGCGCGGTG	2220
Db	2161	TCCATGCGAGAACCTGCTTTTAAAGCGAAACAGCGATGACCATCAACGCGCGGTG	2220
QY	2221	GCATGGTCCAAATCGACACCGCAAGGTAAATTAAGGTGTAACGGTACGCGACCTTC	2280
Db	2221	GCATGGTCCAAATCGACACCGCAAGGTAAATTAAGGTGTAACGGTACGCGACCTTC	2280
QY	2281	AAGGGTCAACCATCCAAACCGTTGCTGGCGTGTGAGCGCAGGTATTAAACGCGCACT	2340
Db	2281	AAGGGTCAACCATCCAAACCGTTGCTGGCGTGTGAGCGCAGGTATTAAACGCGCACT	2340
QY	2341	CCGAAACAAAGAGCTGCAAAAGGTTCTCGAAACTATCTGCTGACTGATGAGGTCTG	2400
Db	2341	CCGAAACAAAGAGCTGCAAAAGGTTCTCGAAACTATCTGCTGACTGATGAGGTCTG	2400
QY	2401	GAGCGGTTAATAAGACAAACCGCTGGGTGCGTAGCGCTGAAAGTCTTACGAGGAAGAG	2460
Db	2401	GAGCGGTTAATAAGACAAACCGCTGGGTGCGTAGCGCTGAAAGTCTTACGAGGAAGAG	2460
QY	2461	TTGGCGAAAGATCCAGTATTCGCGCCACCATGGAACCGCCAGAAAGGTGAAATCATG	2520
Db	2461	TTGGCGAAAGATCCAGTATTCGCGCCACCATGGAACCGCCAGAAAGGTGAAATCATG	2520
QY	2521	CCGAAACATCCCGCAGATGCTCCGCTTCTGCTGATGCGGTGCTGCTGCTGATCAACGCC	2580
Db	2521	CCGAAACATCCCGCAGATGCTCCGCTTCTGCTGATGCGGTGCTGCTGCTGATCAACGCC	2580
QY	2581	GCAGCGGTGCTGAGACTGCTGATGAAAGCCCTGAAAGACGCGCAGCTAAATTCGAGCTCG	2640

Db 2581 GCCAGCGTGTGACAGTGTGATGAAGCCCTGAAAGACGGCGACACTAAATTCGAGCTCG 2640
Qy 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGAGGATTCAGAAATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACCTCGGGATCGAGGAGGATTCAGAAATTC 2700
RESULT 12
LOCUS AX377532 9191 bp DNA linear PAT 18-MAR-2002
DEFINITION Sequence 9 from Patent WO0212553.
ACCESSION AX377532
VERSION AX377532.1 GI:19573718
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1
AUTHORS Kappel,A., Polakowski,T., Pignot,M., Windhab,N., Behrensdoerf,H. and Muth,J.
TITLE Method for detecting mutations in nucleotide sequences
JOURNAL Patent: WO 0212553-A 9 14-FEB-2002;
Nanogen Recognomics GmbH (DE)
FEATURES
source Location/Qualifiers
1..9191
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="pMALc2x-mutS Plaemid"
ORIGIN
Query Match 81.8%; Score 2698; DB 6; Length 9191;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2701; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 CCGACACCATCAATGGTGCAAAACCTTTCGGGGTATGGCATGATAGGCCCGGAAGAGA 60
Db 1 CCGACACCATCAATGGTGCAAAACCTTTCGGGGTATGGCATGATAGGCCCGGAAGAGA 60
Qy 61 GTCAATTACAGGTGGTGAATGTGAACACAGTAACGTTATACCATGTGCGCAGAGTATGCCG 120
Db 61 GTCAATTACAGGTGGTGAATGTGAACACAGTAACGTTATACCATGTGCGCAGAGTATGCCG 120
Qy 121 GTGTCTCTTATCAGACCGGTTTCCTCCGGTGGTGAACACAGCCAGCCACGCTTTCCTCGAAAA 180
Db 121 GTGTCTCTTATCAGACCGGTTTCCTCCGGTGGTGAACACAGCCAGCCACGCTTTCCTCGAAAA 180
Qy 181 CCGGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTTACATTCCTCCAAACCGCGTGGCAC 240
Db 181 CCGGGAAAAAGTGAAGCGCGGATGGCGAGCTGAATTTACATTCCTCCAAACCGCGTGGCAC 240
Qy 241 AACAACTGGCGGCAACAGTCTGCTGATTTGGCGTGGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAACAGTCTGCTGATTTGGCGTGGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 AGCGCGCTGCGCAAAATTTGTCGGGGGATTAATAATCTCGGCCGATCAATCGGTCGACGG 360
Db 301 AGCGCGCTGCGCAAAATTTGTCGGGGGATTAATAATCTCGGCCGATCAATCGGTCGACGG 360
Qy 361 TGGTGTGTCGATGTGTAGAACGAAAGCGCGTGAAGCCCTGTAAGCGCGGTGCACAATC 420
Db 361 TGGTGTGTCGATGTGTAGAACGAAAGCGCGTGAAGCCCTGTAAGCGCGGTGCACAATC 420
Qy 421 TTCTCGGCAACCGCTCAGTGGGTGATCAATTAATCTCCGCTGGATGACAGGATGCCA 480
Db 421 TTCTCGGCAACCGCTCAGTGGGTGATCAATTAATCTCCGCTGGATGACAGGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCTGACATAATGTTCCGGCGTTATTCTTGATGTTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCTGACATAATGTTCCGGCGTTATTCTTGATGTTCTGACCCAGA 540
Qy 541 CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600

Db 541 CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTGTGATTTGGGTACACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Db 601 TGGTGTGATTTGGGTACACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Qy 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Qy 721 CGGAACGGGAAGGAGCTGGAGTGCATGTCGGGTTTTCAACAACCAATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGAGCTGGAGTGCATGTCGGGTTTTCAACAACCAATGCAAAATGCTGA 780
Qy 781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGCCAACGATCAGATGCGCTGGGCGCAA 840
Db 781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGCCAACGATCAGATGCGCTGGGCGCAA 840
Qy 841 TGGCGGCATTTACCGAGTCCGGGCTGGCGTGGTGGCGGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGGCATTTACCGAGTCCGGGCTGGCGTGGTGGCGGATATCTCGGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAAGGATTTTC 960
Qy 961 GCCTGTCTGGGGCAAAACCAAGCGTGGACCGCTTGTGTCGCAACTCTCTCAGGCGCCAGGCGTGA 1020
Db 961 GCCTGTCTGGGGCAAAACCAAGCGTGGACCGCTTGTGTCGCAACTCTCTCAGGCGCCAGGCGTGA 1020
Qy 1021 AGGGCAATCAGCTGTGTCCTCTCATCTGTTGTAAGAAAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGTCCTCTCATCTGTTGTAAGAAAAACCAACCTGGCGCCCAATA 1080
Qy 1081 CGAAACCGCTCTCCCGCGCTTGGCGGATTAATTAATGAGTGGCAGCAGAGTTT 1140
Db 1081 CGAAACCGCTCTCCCGCGCTTGGCGGATTAATTAATGAGTGGCAGCAGAGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGTGAGTTAGTCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGTGAGTTAGTCTCACTCATTTAG 1200
Qy 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACCGGTGACCAATGTTCTGCGG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGACCGGTGACCAATGTTCTGCGG 1260
Qy 1261 TCAGGCAGCCATCGGAAGCTGTGGTATGAGTGTGAGGTGCTGATTAATCACTGATTAATTCG 1320
Db 1261 TCAGGCAGCCATCGGAAGCTGTGGTATGAGTGTGAGGTGCTGATTAATCACTGATTAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACCTCCCGTTCTGGATTAATGTTTTTGGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACCTCCCGTTCTGGATTAATGTTTTTGGCGCGACATCAATACGGTT 1380
Qy 1381 CTGGCAAAATTTCTGAAATGAGCTGTGCAAAATTAATCATCGGCTCGTATTAATGTGTGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTGCAAAATTAATCATCGGCTCGTATTAATGTGTGA 1440
Qy 1441 ATTGTGAGCGGATTAACAAATTTTCAACAGGAAACGCGAGTCCGTTAGGTTGTTTTCAGCA 1500
Db 1441 ATTGTGAGCGGATTAACAAATTTTCAACAGGAAACGCGAGTCCGTTAGGTTGTTTTCAGCA 1500
Qy 1501 GCATTTCAACCAAGGACCATAGATTAATGAAGAGGTAAGGTAATCTGATATCTTGG 1560
Db 1501 GCATTTCAACCAAGGACCATAGATTAATGAAGAGGTAAGGTAATCTGATATCTTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTGGTGAAGAAATTCAGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTGGTGAAGAAATTCAGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTGGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT 1680

Db 1621 ACCGAAATTAAAGTCACCGTTGAGCATCCGATAAACTGGAAGAGAAATTCACACAGGTT 1680
Qy 1681 GCGCAACTGGCGATGCCCTGACATTATCTCTGGGCACACACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGCCCTGACATTATCTCTGGGCACACACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTTGGCTGGAATCACCCCGGACAAAGCGTTCCAGGACAACTGTAT 1800
Db 1741 GCTCAATCTGGCTGTTGGCTGGAATCACCCCGGACAAAGCGTTCCAGGACAACTGTAT 1800
Qy 1801 CCGTTTACCTGGGATCGCTACGTTTACAGCGCAAGCTGATTCCTTACCGATTCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATCGCTACGTTTACAGCGCAAGCTGATTCCTTACCGATTCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATAACAAGATCTCTCCGAACCCGCCAAAACCTCGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAACAAGATCTCTCCGAACCCGCCAAAACCTCGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTCAAGAACCGTACTTTCACCTGGCGCTGATTCGCTGACGCGGGTTATCGGTTCAAG 2040
Db 1981 CTCAAGAACCGTACTTTCACCTGGCGCTGATTCGCTGACGCGGGTTATCGGTTCAAG 2040
Qy 2041 TATGAAACGGCAAGTACGACATTAAGACGTGGCGCTGCATAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAACGGCAAGTACGACATTAAGACGTGGCGCTGCATAACGCTGGCGCGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGTTGACCTGATTAAGAACTGAAGCGTGGCGTGGATAACGCTGGCGCGAAAGCG 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAAGAACTGAAGCGTGGCGTGGATAACGCTGGCGCGAAAGCG 2160
Qy 2161 TCCATCGCAAGAGCTGCTCTTAAATAAGCGCAACAGCGATGACCATCAAGCGCCCGTGG 2220
Db 2161 TCCATCGCAAGAGCTGCTCTTAAATAAGCGCAACAGCGATGACCATCAAGCGCCCGTGG 2220
Qy 2221 GCATGCTCAACATCGACACAGCAAGAGTAAATATGTTGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACAGCAAGAGTAAATATGTTGTAACGGTACTGCGGACCTTC 2280
Qy 2281 AAGGGTCAACCATCAAAACCGTTCTGCGCTGCTGAGCGAGGTATTAAGCGCCCGT 2340
Db 2281 AAGGGTCAACCATCAAAACCGTTCTGCGCTGCTGAGCGAGGTATTAAGCGCCCGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTTCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTCGGTGGCTGAGCTGAAAGTCTTACGAGAAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTCGGTGGCTGAGCTGAAAGTCTTACGAGAAAG 2460
Qy 2461 TTGGCGAAAGATCCACGTTATGCGCCACCATGGAACCGCCGAGAAAGGTTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATGCGCCACCATGGAACCGCCGAGAAAGGTTGAATCATG 2520
Qy 2521 CCGAACATCCGCGAGATGTCGCTTTCTGGTATGCCGCTGCTGAGCTGATCAAGCGCC 2580
Db 2521 CCGAACATCCGCGAGATGTCGCTTTCTGGTATGCCGCTGCTGAGCTGATCAAGCGCC 2580
Qy 2581 GCCAGCGTCTGACGCTGATGAAGCGCTGAAAGCGCGGAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACGCTGATGAAGCGCTGAAAGCGCGGAGACTAATTCGAGCTCG 2640
Qy 2641 AACAAACAAACAATAACAATAACAACACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACAAACAAACAATAACAATAACAACACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Qy 2701 GGATCC 2706
Db 2701 GGATCC 2706

RESULT 13
AF097412
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. .6724
/organism="Expression vector pMBP-parallel1"
/mol_type="other DNA"
/db_xref="taxon:84479"
204. .1163
/gene="lacIq"
204. .1163
/gene="lacIq"
/codon_start=1
/transl_table=11
/product="lac repressor"
/protein_id="AAC83812.1"
/db_xref="GI:3983121"
/translation="MAELNYIPNRVAQAGKQSLIGVATSSLAHPQSQIVAAIKS
RADQLGASVVSVMVERSGVEACAAVHNLIAQVSLIINYPLDDODAIAVEAACTNV
PALFLDSVQTLPTINSIFSHEDGTRGLVHVALGHQOIALLAGPLSSVSRILAGW
HKYLTNRQIPTAERGDWSMGFOOTMQLNEGIVPTAMLVANDQMALGAMRAITE
SGLRVGADI.SVVYDDTDSVCYIPPLTTIKODRLLGQTS.DRLQLQLSQGAQVKGNG
LIPVSLVKEKTLIAENTQTASPRALADSLMARQVSRLESQ"
1453. .1481
/note="tac; inducible promoter"
1528. .2994
/gene="male"
1528. .2994
/gene="male"
/codon_start=1
/transl_table=11
/product="maltose binding protein"
/protein_id="AAC83813.1"
/db_xref="GI:3983122"
/translation="MKIEGKLVIVINGDKYNGLAIEVGKKFEKDTGKVTVEHPDKL
ESKFPVAATGDPDIIFWHDRCGYAQSGLLAIEITDKAFQDKLYEFTWDVARYNG
KLIAVPIAEVALSLIYNKDLLNPETWBEIPALDKELKAKGSKALMNFLOEYFTWP
LIAADGGTAFYENGKDYDKDVGVDNAGAKAGLTFLVDLIKNKHNDADTYSIAEAA
NKGETAMTINGPWANSNDTSKVNYGVTVLPTFKGQSPFVGLSAGTNAASPNKEL
AKCEFLNYLLTDDEGLEAVNKDKPLGAVALKSYEEELAKDPRIATAEMNAQKGEIMPN
PQMSAFVAVRTAVNAASGRTVDALKDAQTSNNSSNNNNNNNNNNI.GIDTTELYP
QGAQPEFGLRRRAQLVRPLSNLEPAVSRHVPVSLALAVLVLRDDWENPGVTQUNRL
AAHPFPASWRNSEARTDRPSQQLRSLNGEWQLGCGFG"
2633. .2688
/gene="male"
/note="protein spacer region"
2689. .2709
/gene="male"
/note="Tobacco etch virus protease cleavage site"
2711. .2810
/gene="male"

AF097412 6724 bp DNA circular SYN 28-JAN-2002
Expression vector pMBP-parallel1, complete sequence.
AF097412
AF097412.1 GI:3983120
Expression vector pMBP-parallel1
Expression vector pMBP-parallel1
other sequences: artificial sequences; vectors.
1 (bases 1 to 6724)
Sheffield, P., Garrard, S. and Derewenda, Z.
Overcoming expression and purification problems of RhoGDI using a
family of 'parallel' expression vectors
Protein Expr. Purif. 15 (1), 34-39 (1999)
99150479
10024467
2 (bases 1 to 6724)
Sheffield, P.J., Garrard, S.M. and Derewenda, Z.S.
Direct Submission
Submitted (05-OCT-1998) Molecular Physiology and Biological
Physics, University of Virginia, 4215 Jordan Hall, 1300 Jefferson
Park Avenue, Charlottesville, Virginia 22908, USA
Location/Qualifiers
1. .6724
/organism="Expression vector pMBP-parallel1"
/mol_type="other DNA"
/db_xref="taxon:84479"
204. .1163
/gene="lacIq"
204. .1163
/gene="lacIq"
/codon_start=1
/transl_table=11
/product="lac repressor"
/protein_id="AAC83812.1"
/db_xref="GI:3983121"
/translation="MAELNYIPNRVAQAGKQSLIGVATSSLAHPQSQIVAAIKS
RADQLGASVVSVMVERSGVEACAAVHNLIAQVSLIINYPLDDODAIAVEAACTNV
PALFLDSVQTLPTINSIFSHEDGTRGLVHVALGHQOIALLAGPLSSVSRILAGW
HKYLTNRQIPTAERGDWSMGFOOTMQLNEGIVPTAMLVANDQMALGAMRAITE
SGLRVGADI.SVVYDDTDSVCYIPPLTTIKODRLLGQTS.DRLQLQLSQGAQVKGNG
LIPVSLVKEKTLIAENTQTASPRALADSLMARQVSRLESQ"
1453. .1481
/note="tac; inducible promoter"
1528. .2994
/gene="male"
1528. .2994
/gene="male"
/codon_start=1
/transl_table=11
/product="maltose binding protein"
/protein_id="AAC83813.1"
/db_xref="GI:3983122"
/translation="MKIEGKLVIVINGDKYNGLAIEVGKKFEKDTGKVTVEHPDKL
ESKFPVAATGDPDIIFWHDRCGYAQSGLLAIEITDKAFQDKLYEFTWDVARYNG
KLIAVPIAEVALSLIYNKDLLNPETWBEIPALDKELKAKGSKALMNFLOEYFTWP
LIAADGGTAFYENGKDYDKDVGVDNAGAKAGLTFLVDLIKNKHNDADTYSIAEAA
NKGETAMTINGPWANSNDTSKVNYGVTVLPTFKGQSPFVGLSAGTNAASPNKEL
AKCEFLNYLLTDDEGLEAVNKDKPLGAVALKSYEEELAKDPRIATAEMNAQKGEIMPN
PQMSAFVAVRTAVNAASGRTVDALKDAQTSNNSSNNNNNNNNNNI.GIDTTELYP
QGAQPEFGLRRRAQLVRPLSNLEPAVSRHVPVSLALAVLVLRDDWENPGVTQUNRL
AAHPFPASWRNSEARTDRPSQQLRSLNGEWQLGCGFG"
2633. .2688
/gene="male"
/note="protein spacer region"
2689. .2709
/gene="male"
/note="Tobacco etch virus protease cleavage site"
2711. .2810
/gene="male"

gene /note="multiple cloning site"
3426..4356
/gene="bla"
promoter /gene="bla"
3426..3454
/gene="bla"
3496..4356
/gene="bla"
CDS /codon_start=1
/transl_table=11
/product="beta-lactamase"
/protein_id="AAC83814.1"
/db_xref="GI:3983123"
/translation="MSIOHFRVALIPFPAFLPFAHPETLVKVKDAEDQLGARVGY
VELDLSNGKILSPFRSPRSNMSTFKVLLCGNAVLSRDVAGQEGLGRRIHYSDNOLVE
YSPVTEKHLTDGMTVRELSAAITWSDNTANALLTIIGGPKBELTALFHNMGDHLRL
DRWPELNEAIPNDREDITMPVAMATTLKLTGELLTLASQQIIDLMEADKVGPL
LRSLPAGWFIADKSGAGBRGSRGIIAALGPDGKPSRIIVITVITGSGQATMDERNQIA
EIGASLIKHW"

ORIGIN

Query Match 81.1%; Score 2676.2; DB 12; Length 6724;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2678; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGCAGAACCTTTCCGGGTATGGCATGATAGCCCGGAGAGA 60
DB 1 CCGACACCATCGAATGGTGCAGAACCTTTCCGGGTATGGCATGATAGCCCGGAGAGA 60
QY 61 GTCAATTACAGGTGGTGAATGTGAACACAGTAACTGATACCATGTCGAGAGTATGCCG 120
DB 61 GTCAATTACAGGTGGTGAATGTGAACACAGTAACTGATACCATGTCGAGAGTATGCCG 120
QY 121 GTGTCCTTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCACGTTTCTGGCAAAA 180
DB 121 GTGTCCTTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCACGTTTCTGGCAAAA 180
QY 181 CCGGGGAAAAAGTGGAGCGGATGGCGAGTGAATTAATTCCTCAACCGCGTGGGCAC 240
DB 181 CCGGGGAAAAAGTGGAGCGGATGGCGAGTGAATTAATTCCTCAACCGCGTGGGCAC 240
QY 241 AACCACTGGCGGCAACAGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
DB 241 AACCACTGGCGGCAACAGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTCGCAAAATTTGTCGGCGGATTTAAATCTCGCGCGATCAACTGGGTGCCAGCG 360
DB 301 ACGCGCGTCGCAAAATTTGTCGGCGGATTTAAATCTCGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGTGTGATGTAGAACGAGCGGCTGGAAGCCTGTAAAGCGGCGTGCACAATC 420
DB 361 TGGTGTGTGATGTAGAACGAGCGGCTGGAAGCCTGTAAAGCGGCGTGCACAATC 420
QY 421 TTCTCGGCAACCGTCAGTGGGCTGATCATTTAACTATCCGCTGGATGACAGGATGCCA 480
DB 421 TTCTCGGCAACCGTCAGTGGGCTGATCATTTAACTATCCGCTGGATGACAGGATGCCA 480
QY 481 TTGCTGTGGAAGTGCCTGCATTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
DB 481 TTGCTGTGGAAGTGCCTGCATTAATGTTCCGGCGTTATTTCTTGATGTCTCTGACCAGA 540
QY 541 CACCCATCAACAGTATTTATTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTTATTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
QY 601 TGGTGCATTTGGTTCACCGCAATCCGCTGTGTAGCGGCCCATTAAGTTCTGTCTCGG 660
DB 601 TGGTGCATTTGGTTCACCGCAATCCGCTGTGTAGCGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CCGGCTCTGGCTCTGGCTGGCTGATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CCGGCTCTGGCTCTGGCTGGCTGATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CCGAACGGGAAGCGCACTGGAGTGCATGTCCGGTTTTCACAAACCATGCAATGCTGTA 780

DB 721 CCGAACGGGAAGCGCACTGGAGTGCATGTCCGGTTTTCACAAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTCGATGCTGGTTGCCAACGATCAGATGCGCTGGCGCGAA 840
DB 781 ATGAGGGCATCGTTCCCACTCGATGCTGGTTGCCAACGATCAGATGCGCTGGCGCGAA 840
QY 841 TCGCGCCATTAACGAGTCCGGGCTGGCGGTGGTGGCGGATATCTCGGTAGTGGGATAG 900
DB 841 TCGCGCCATTAACGAGTCCGGGCTGGCGGTGGTGGCGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACACCATCAAAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTAAACACCATCAAAACAGGATTTTC 960
QY 961 GCCTCTCGGGGCAAAACACGAGCTGGAACCGCTTCTCTCAGGCGCAGCGGTGA 1020
DB 961 GCCTCTCGGGGCAAAACACGAGCTGGAACCGCTTCTCTCAGGCGCAGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGGCCCGTCTCAGCTGTTGAAAGAAAAACCACTGGCGCCCAATA 1080
DB 1021 AGGGCAATCAGCTGTGGCCCGTCTCAGCTGTTGAAAGAAAAACCACTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTGGCGGATTCATTAATGCAGCTGCACACAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGGTGGCGGATTCATTAATGCAGCTGCACACAGGTTT 1140
QY 1141 CCGGACTCGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
DB 1141 CCGGACTCGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAATTTCTCATGTTTGCAGCTTATCATCGACTGCAACGCTGCACTGTTCTGGCG 1260
DB 1201 GCACAATTTCTCATGTTTGCAGCTTATCATCGACTGCAACGCTGCACTGTTCTGGCG 1260
QY 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGTGCAGCTGTAATCACTGATTAATTCG 1320
DB 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGTGCAGCTGTAATCACTGATTAATTCG 1320
QY 1321 TGTCCCTCAAGCGCACCTCCCGTCTCGGATAATGTTTTTGGCGCGACATCAATAACGGTT 1380
DB 1321 TGTCCCTCAAGCGCACCTCCCGTCTCGGATAATGTTTTTGGCGCGACATCAATAACGGTT 1380
QY 1381 CTGGCAATAATTTCTGAAATGAGTCTGTGCAATTAATCACTCGCTCGTATAATGTGTGA 1440
DB 1381 CTGGCAATAATTTCTGAAATGAGTCTGTGCAATTAATCACTCGCTCGTATAATGTGTGA 1440
QY 1441 ATTGTGAGCGGATAACAATTTTCAACAGGAACACCGCTGCGTTAGTGTGTTTTCACGA 1500
DB 1441 ATTGTGAGCGGATAACAATTTTCAACAGGAACACCGCTGCGTTAGTGTGTTTTCACGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACTGAAGAAGGTAAACTGGTAATCTGG 1560
DB 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACTGAAGAAGGTAAACTGGTAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACCGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTAACGGCGATAAAGGCTATAACCGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGAATAACTGGAAGAGAAATTCCTCCACAGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGAATAACTGGAAGAGAAATTCCTCCACAGTT 1680
QY 1681 GGGGCACTGGCGATGGCCCTGACATTTCTTCTGGGACACGCGCTTTCGTGGCTAC 1740
DB 1681 GGGGCACTGGCGATGGCCCTGACATTTCTTCTGGGACACGCGCTTTCGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTGGCTGAAATCAACCCCGGCAAGCGTTCCAGSACAAGCTGAT 1800
DB 1741 GCTCAATCTGGCCTGTGGCTGAAATCAACCCCGGCAAGCGTTCCAGSACAAGCTGAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTAGCTTACACGGCAAGCTGATTTGCTTACCCGATCCGTGTT 1860
DB 1801 CCGTTTACCTGGGATGCGGTAGCTTACACGGCAAGCTGATTTGCTTACCCGATCCGTGTT 1860

Db 1801 CCGTTTACCTGGGATGCGCTACGTTACAAACGCGAAGCTGATTGCTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTCATTTATAACAAGATCTGTCGCGAACCGCCCAAAACCTGGAA 1920
Db 1861 GAAGCGTTATCGCTCATTTATAACAAGATCTGTCGCGAACCGCCCAAAACCTGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGAGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGAGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGGCAAGTACGACATTAAAGAGCGTGGCGTGATAAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGGCAAGTACGACATTAAAGAGCGTGGCGTGATAAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCCCTGGTTGACCTGATTAAACAAACACATGAATGACACCGGATTAC 2160
Db 2101 GGTCTGACCTTCCCTGGTTGACCTGATTAAACAAACACATGAATGACACCGGATTAC 2160
QY 2161 TCCATCGCAGAAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACCGCGCGTGG 2220
Db 2161 TCCATCGCAGAAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACCGCGCGTGG 2220
QY 2221 GCATGTCGAACATCGACACCGAAGTAAATGATGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGTCGAACATCGACACCGAAGTAAATGATGTAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGTCGAACATCCCAACCGTTTCGTTGGCGTCTGAGCGCGAGTATTAAACCGCGCGAGT 2340
Db 2281 AAGGTCGAACATCCCAACCGTTTCGTTGGCGTCTGAGCGCGAGTATTAAACCGCGCGAGT 2340
QY 2341 CCAGAACAAAGAGTGGCAAAAGAGTTCTCGAAACCTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCAGAACAAAGAGTGGCAAAAGAGTTCTCGAAACCTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACAGGAAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACAGGAAAG 2460
QY 2461 TTGGCGAAAGATCCAGTATTGCGCGCACCATGGAAGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTGCGCGCACCATGGAAGAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACGCTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACGCTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACAGTCTGATGAAGCCCTGGAAGACGGCGAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTCTGATGAAGCCCTGGAAGACGGCGAGACTAATTCGAGCTCG 2640
QY 2641 AACACAAACAATAACAATAACAACACCTCGGGATCGA 2681
Db 2641 AACACAAACAATAACAATAACAACACCTCGGGATCGA 2681

RESULT 14
AX378208
LOCUS
DEFINITION Sequence 194 from Patent W00206306.
ACCESSION AX378208
VERSION
KEYWORDS AX378208.1 GI:19574061
SOURCE unidentified
ORGANISM unclassified.
REFERENCE 1
AUTHORS Yan, R., Tomasselli, A.G., Gurney, M.E., Emmons, T.L., Bienkowski, M.J.
and Heinrikson, R.L.
TITLE Substrates and assays for _g(b)-secretase activity

JOURNAL Patent: WO 0206306-A 194 24-JAN-2002;
FEATURES PHARMACIA & UPJOHN COMPANY (US)
source Location/Qualifiers
1. 6806
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
/note="fusion protein comprising a maltose binding protein
with 125 amino"

Query Match 79.9%; Score 2636.8; DB 6; Length 6806;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2638; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

ORIGIN

QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTTCGCGGTATGGCATGATAGCCGCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAGAAACCTTTTCGCGGTATGGCATGATAGCCGCCGGAAGAGA 60
QY 61 GTCATTTCCAGGTTGTAATGTGAACACAGTAACGTTTATACGATGTCCGAGAGTATGCCG 120
Db 61 GTCATTTCCAGGTTGTAATGTGAACACAGTAACGTTTATACGATGTCCGAGAGTATGCCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCCGCTGTGTGAACAGCCAGCCAGCTTTCTCGGAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGCTGTGTGAACAGCCAGCCAGCTTTCTCGGAAA 180
QY 181 CCGCGGAAAAGTGAAGCGCGATGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
Db 181 CCGCGGAAAAGTGAAGCGCGATGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGCAACAGTCGTTGATTTGGCGTTGCGCCTGACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAACAGTCGTTGATTTGGCGTTGCGCCTGACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTCCGAAATTTGTTCGCGCGATTAATTTCTCGCGCGATTAATTTCTCGCGCGAT 360
Db 301 ACGCGCGTCCGAAATTTGTTCGCGCGATTAATTTCTCGCGCGATTAATTTCTCGCGCGAT 360
QY 361 TGTGTGTGTCGATGTTAGAACGAGCGGTGTAAGCGCGTGTAAAGCGCGGTGCAAAATC 420
Db 361 TGTGTGTGTCGATGTTAGAACGAGCGGTGTAAGCGCGTGTAAAGCGCGGTGCAAAATC 420
QY 421 TTCTCGCGCAACCGTCACTGTTGCTGATCAATTAATTTCTCGCGCGATTAATTTCTCGCGCG 480
Db 421 TTCTCGCGCAACCGTCACTGTTGCTGATCAATTAATTTCTCGCGCGATTAATTTCTCGCGCG 480
QY 481 TTGCTGTGAGCTGCTGCACTAATGTTCCGCGCTTATTTCTTGATGTCTTGACCAGA 540
Db 481 TTGCTGTGAGCTGCTGCACTAATGTTCCGCGCTTATTTCTTGATGTCTTGACCAGA 540
QY 541 CACCCATCAACAGTATTTTCTCCCATGAAGCGGTACGCGACTGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTTCTCCCATGAAGCGGTACGCGACTGGCGTGGAGCATC 600
QY 601 TGTGTCGATTTGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCGG 660
Db 601 TGTGTCGATTTGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCGG 660
QY 661 CCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 661 CCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 CCGAAACGGGAAGCGGACCTGAGTGCATGTCGCGTTTTCACAAACCAACATGAAATGCTGA 780
Db 721 CCGAAACGGGAAGCGGACCTGAGTGCATGTCGCGTTTTCACAAACCAACATGAAATGCTGA 780
QY 781 ATGAGGCGCATCGTTCCGACTGCGATGCTGTTGCCAAGATGATGCGCTGGCGCAA 840
Db 781 ATGAGGCGCATCGTTCCGACTGCGATGCTGTTGCCAAGATGATGCGCTGGCGCAA 840
QY 841 TCGCGCGCATTAACCGAGTCCGCGTGGTGGCGGATATCTCGGTATGCGGATACG 900
Db 841 TCGCGCGCATTAACCGAGTCCGCGTGGTGGCGGATATCTCGGTATGCGGATACG 900

841 TCGCGCCATTACCGAGTCCGGGCTCGGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACCAACCATCAAAACAGGATTTTC 960
901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACCAACCATCAAAACAGGATTTTC 960
961 GCCTGCTGGGGCAACACGAGCTGGAGCGGTCTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
961 GCCTGCTGGGGCAACACGAGCTGGAGCGGTCTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTCAAAAGAAACACCACTGGCGGCCCAATA 1080
1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTCAAAAGAAACACCACTGGCGGCCCAATA 1080
1081 GCGAAACCGCCTCTCCCGCGGTTGGCGGATTCATTAAATGCGAGCTGGCAGCAGGTTT 1140
1081 GCGAAACCGCCTCTCCCGCGGTTGGCGGATTCATTAAATGCGAGCTGGCAGCAGGTTT 1140
1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGCGGTGCAACCAATGCTTCTGGCG 1260
1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGCGGTGCAACCAATGCTTCTGGCG 1260
1261 TCAGGAGCGCATCGGAAGCTGGTATGCGTGTGAGGTGCGTAATCACTGATCAATTCG 1320
1261 TCAGGAGCGCATCGGAAGCTGGTATGCGTGTGAGGTGCGTAATCACTGATCAATTCG 1320
1321 TGTGCTCAAGCGGCACTCCCGTCTGATATATGTTTTTGGCGCGACATCAATACGGTT 1380
1321 TGTGCTCAAGCGGCACTCCCGTCTGATATATGTTTTTGGCGCGACATCAATACGGTT 1380
1381 CTGGCAATATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGA 1440
1381 CTGGCAATATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGA 1440
1441 ATTGTGAGCGGATAACAAATTTACACAGGAAACAGCCAGTCGGTTTAGTGTTTTTCAGGA 1500
1441 ATTGTGAGCGGATAACAAATTTACACAGGAAACAGCCAGTCGGTTTAGTGTTTTTCAGGA 1500
1501 GCACCTCACCAACAGGACCATAGATTTATGAAAACCTGAAGAAGGTAAACTGTTGTAATCTGG 1560
1501 GCACCTCACCAACAGGACCATAGATTTATGAAAACCTGAAGAAGGTAAACTGTTGTAATCTGG 1560
1561 ATTAACGGCGATAAAGGCTATAACCGTCTCGCTGAAGTTCGGTAAGAAATTCGAGAAAGAT 1620
1561 ATTAACGGCGATAAAGGCTATAACCGTCTCGCTGAAGTTCGGTAAGAAATTCGAGAAAGAT 1620
1621 ACCGGAATTTAAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCACAGGTT 1680
1621 ACCGGAATTTAAAGTCACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCACAGGTT 1680
1681 GCGGCAATCGGCGATGGCCCTGACATTTATCTTGGGCAACAGCGCTTTGGTGGCTAC 1740
1681 GCGGCAATCGGCGATGGCCCTGACATTTATCTTCTGGGCAACAGCGCTTTGGTGGCTAC 1740
1741 GCTCAATCTGGCGCTGTGGCTGAAATCAACCCCGCAAGCGTTCCAGGACAAGCTGTAT 1800
1741 GCTCAATCTGGCGCTGTGGCTGAAATCAACCCCGCAAGCGTTCCAGGACAAGCTGTAT 1800
1801 CCGTTTACCTGGGATGCGGTACGTTTACACGCGCAAGCTGATTTACCCGATCGCTGTT 1860
1801 CCGTTTACCTGGGATGCGGTACGTTTACACGCGCAAGCTGATTTACCCGATCGCTGTT 1860
1861 GAAGCGTTATCGCTGATTTATCAAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
1861 GAAGCGTTATCGCTGATTTATCAAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA 1920
1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980

QY 1981 CTGCAAGAACCGTACTTTCACCTGGCCGCTGATTCTGCTGACGGGGTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGGCCGCTGATTCTGCTGACGGGGTTATCGGTTCAAG 2040
QY 2041 TATGAAAACGCGAAGTACGACATTAAAGACCTGGCGGTGGATTAACGCTGGCGCGAAGCGG 2100
Db 2041 TATGAAAACGCGAAGTACGACATTAAAGACCTGGCGGTGGATTAACGCTGGCGCGAAGCGG 2100
QY 2101 GGTCTGACCTTCTGGTTGACCTGATTAAAAACAAACACATGAATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAAAAACAAACACATGAATGACAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAAGCTGCTTTAATAAAGCCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAAGCTGCTTTAATAAAGCCGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCACATCGACACGACCAAGTGAATTAATGGTGTAAACGCTACTCCGACCTTC 2280
Db 2221 GCATGGTCCACATCGACACGACCAAGTGAATTAATGGTGTAAACGCTACTCCGACCTTC 2280
QY 2281 AAGGTCACACCATCCAAACCGTTCGTTGGCGTGTGAGCGAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTCACACCATCCAAACCGTTCGTTGGCGTGTGAGCGAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
QY 2401 GAAGGGTTAATAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
Db 2401 GAAGGGTTAATAAGACAAACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
QY 2461 TTGGCAAAAGATCCACGTTATTTGCCGCCACACCTGAGAAAAACGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGCAAAAGATCCACGTTATTTGCCGCCACACCTGAGAAAAACGCCAGAAAGGTGAAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGGTTTTCTGGTATGCCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2580
Db 2521 CCGAACATCCCGCAGATGTCGGTTTTCTGGTATGCCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2580
QY 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640

RESULT 15
AF031088 6706 bp DNA circular SYN 07-MAY-2001
LOCUS Shuttle vector pMAL-pIII, complete sequence.
DEFINITION AF031088
ACCESSION AF031088
VERSION AF031088.1 GI:2623821
KEYWORDS Shuttle vector pMAL-pIII
SOURCE Shuttle vector pMAL-pIII
ORGANISM Shuttle vector pMAL-pIII
OTHER SEQUENCES; artificial sequences; vectors.
REFERENCE 1 (bases 1 to 6706)
AUTHORS Zwick,M.B., Bonnycastle,L.L., Noren,K.A., Venturini,S., Leong,E.,
Barbas,C.F. III, Noren,C.J. and Scott,J.K.
TITLE The maltose-binding protein as a scaffold for monovalent display of
peptides derived from phage libraries
JOURNAL Anal. Biochem. 264 (1), 87-97 (1998)
MEDLINE 99002881
PUBMED 9784192
REFERENCE 2 (bases 1 to 6706)
AUTHORS Noren,K.A. and Noren,C.J.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-1997) New England Biolabs, 32 Tozer Road,
Beverly, MA 01915, USA
FEATURES location/Qualifiers
source 1..6706
/organism="Shuttle vector pMAL-pIII"
/mol_type="genomic DNA"


```
/db_xref="taxon:69327"
/notes="derived from commercially available vector pMal-p2
(New England Biolabs)"
15..1163
/gene="lacI"
/promoter
/notes="E. coli gene"
15..46
/gene="lacI"
81..1163
/gene="lacI"
/notes="transcriptional repressor"
/codon_start=1
/transl_table=11
/product="lac repressor"
/protein_id="AAB86558.1"
/db_xref="GI:2623822"
/translation="MKPVLVDVAYAGVSQYTVSRVNVQASHVSAKTRKVEAAMAE
LVYINRVAQAGKQSLIGVTSLSALHAPSQIVAAIKSRADQLGASVVSVMVER
GVKCAKAAVHNLAKQVSLIINYPLDDQDAIAVEAACTNVPALFLVSDQTPINSII
FSHEDTRLGVHVALGHQIALLAGPLSSVARSRLAGWHKYLTRNQIOPIABREG
DMSAGSFQOTQMINEGIVPTAMLVANDQMALGAMRAITSGSLRGADISVVGDDT
EDSSCYIPPLTTIKODPRLLGOTSVDRLQLSQQAVKQQLLVSLVKRKTTLAPNT
QTASPRALADSLMLQARQVSRLESQ"
1406..1433
/promoter
/notes="Pac promoter"
1528..2976
/gene="male-lacZa fusion"
/notes="E. coli gene fusion"
1528..2976
/gene="male-lacZa fusion"
/notes="MBP-lacZa; MBP modified to include phage M13
secretion signal sequence"
/codon_start=1
/transl_table=11
/product="maltose binding protein-lacZ alpha peptide
fusion protein precursor"
/protein_id="AAB86559.1"
/db_xref="GI:2623823"
/translation="MKKLLFAPLVLPVFFSHSADMKBEGKLVIIWINGDKYNGLAEV
GKFKEDQYIKVTVHPDKLEKFPQVATGDGPDIIFWAHDHFGVYQSGLLAEITP
DAQFDQILPPTWADRVNKGILAVIAVEALSILYKDLIPNPKTWEIIPALDKEL
KAKGSAALMFLNOBEPYFWPLIADGGVAFKVGKDYIDKVDVGNAGAKAGLFLVD
LIFKNHMADTDYSIAEAFNKGETAMTINGPWSNIDTSKNGYGVTVLPFKGQPS
KPFVGLSAGINAASPNKELAKFLEFILLTDEGLEAVNKDKPLGVALKSYEBELAK
DPRIATNENAKGIMPNIPQMSAFWAVETAIVNASGRQTVDEALKDAQTNSSN
NNNNNNNLGIESRISFSSRVDLQASLALAVLVLRDWNENPGVTQLNRLAAHPFP
ASWRNSEARTDRPSQQLRSLNGEWQLGCFG"
1528..1581
/gene="male-lacZa fusion"
/notes="M13 secretion signal sequence"
1582..2973
/gene="male-lacZa fusion"
/product="maltose binding protein-lacZ alpha peptide
fusion protein"
3478..4338
/gene="bla"
/notes="R-plasmid ampicillin resistance gene"
3478..4338
/gene="bla"
/notes="penicillinase"
/codon_start=1
/transl_table=11
/product="pTEM-1 beta-lactamase"
/protein_id="AAB86560.1"
/db_xref="GI:2623824"
/translation="MSIQHFRVALIPFFAAFLPVFAHPETLVKVKDAEDQLGARVGY
IELDLSKKILESFRPEKPMSTFKVLLCGAVLSRVDAQOLGRRIHYSNDLVE
YSPVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTIGGPKELTAFLEHMGPHVTEL
DRWPELNEATLPNDRDTPWVAATTLKLLTGELTLASRQLDLDMWADKVAQPL
LRSLALPAGWFTADSKAGRGSRGIILALGPDGKPSRIVVIYITGSAQTMDERNRQIA
EIGSLIKHW"
4381..4893
/standard_name="M13 origin (+ -)"
```

```
ORIGIN
Query Match
Best Local Similarity 79.7%; Score 2631; DB 12; Length 6706;
Matches 2704; Conservative 0; Mismatches 5; Indels 60; Gaps 1;
QY 1 CGGACACCATCGAATGGTGTGCAAAACCTTTTCGCGTATGGCATGATAGCCCGGAAGAGA 60
DB 1 CGGACACCATCGAATGGTGTGCAAAACCTTTTCGCGTATGGCATGATAGCCCGGAAGAGA 60
QY 61 GTCGAATTCAGGCGTGTGAATGTGAACCAAGTAACTTATACGATGTGCAGAGTATGCCG 120
DB 61 GTCGAATTCAGGCGTGTGAATGTGAACCAAGTAACTTATACGATGTGCAGAGTATGCCG 120
QY 121 GTGCTCTTATCAGACCGTGTTCGCGGTGTGAACCAAGTAACTTATACGATGTGCAGAGTATGCCG 180
DB 121 GTGCTCTTATCAGACCGTGTTCGCGGTGTGAACCAAGTAACTTATACGATGTGCAGAGTATGCCG 180
QY 181 CGCGGGAAGAGTGAAGCGCGGATGGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
DB 181 CGCGGGAAGAGTGAAGCGCGGATGGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACTCCAGTCTGGCCCTGC 300
QY 301 ACGCCGCGTGCAGAAATTTGTCCGCGGATTAATTCCTCCGCGCGATCAACTGGGTGCCACG 360
DB 301 ACGCCGCGTGCAGAAATTTGTCCGCGGATTAATTCCTCCGCGCGATCAACTGGGTGCCACG 360
QY 361 TGGTGGTGTGATGTGAAGCAAGCGGCTCGAAGCTGTAAAGCGCGGTGACACATC 420
DB 361 TGGTGGTGTGATGTGAAGCAAGCGGCTCGAAGCTGTAAAGCGCGGTGACACATC 420
QY 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCAATTAATTCCTCCGCTGGATGACACGATGCCA 480
DB 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCAATTAATTCCTCCGCTGGATGACACGATGCCA 480
QY 481 TTGCTGTGGAGCTGCTGCACCTAATGTTCCGCGGTTATTTCTTGATGTTCTTGACAGA 540
DB 481 TTGCTGTGGAGCTGCTGCACCTAATGTTCCGCGGTTATTTCTTGATGTTCTTGACAGA 540
QY 541 CACCCATCAACAGTATTTTCTCCATGAACACGCTAGCGACTGGCGTGGCGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTTTCTCCATGAACACGCTAGCGACTGGCGTGGCGTGGAGCATC 600
QY 601 TGGTGGATTTGGGTCAACAGCAAAATCGCGTGTAGCGGGCCCATTAAGTTCTGTCTCG 660
DB 601 TGGTGGATTTGGGTCAACAGCAAAATCGCGTGTAGCGGGCCCATTAAGTTCTGTCTCG 660
QY 661 CGCGTCTGGCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGGCTGCGTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGCGCACTGGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTCA 780
DB 721 CGGAACGGGAAGCGCACTGGAGTGCCATGTCCGGTTTTTCAACAAACCATGCAAAATGCTCA 780
QY 781 ATGAGGGCATCGTTCCCACTCGGATGCTGTTGCCAACGATCAGATGGCGTGGGCGCAA 840
DB 781 ATGAGGGCATCGTTCCCACTCGGATGCTGTTGCCAACGATCAGATGGCGTGGGCGCAA 840
QY 841 TGGCGGCCATTACCGAGTCCGGCTGCGGCTGGTGGCGATATCTCGGTAGTGGGATAG 900
DB 841 TGGCGGCCATTACCGAGTCCGGCTGCGGCTGGTGGCGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAAAACAGGATTTTC 960
DB 901 ACGATACCAAGACAGCTCATGTTATATCCCGCGTTAAACCAATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTCTCCTCAGGGCCAGGCGGTGA 1020
DB 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTCTCCTCAGGGCCAGGCGGTGA 1020
```

Qy	1021	AGGCAATCAAGCTGTTGCCGTCCTCACTGGTGAAGAAAAACACCTTGGCGCCCAATA	1080
Db	1021	AGGCAATCAAGCTGTTGCCGTCCTCACTGGTGAAGAAAAACACCTTGGCGCCCAATA	1080
Qy	1081	CGAAACCGCTCTCCCGCGGCTTGGCCGATTCATTAAATGCAGCTGGCAGCAGAGTTT	1140
Db	1081	CGAAACCGCTCTCCCGCGGCTTGGCCGATTCATTAAATGCAGCTGGCAGCAGAGTTT	1140
Qy	1141	CCGACTGGAAGCGGCAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCAATTAG	1200
Db	1141	CCGACTGGAAGCGGCAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCAATTAG	1200
Qy	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTTGGCG	1260
Db	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTTGGCG	1260
Qy	1261	TCAGCAGCCCATCGAAAGCTGTGTATGGCTGTGCAGGTGCTAAATCACTGCATAATTCG	1320
Db	1261	TCAGCAGCCCATCGAAAGCTGTGTATGGCTGTGCAGGTGCTAAATCACTGCATAATTCG	1320
Qy	1321	TGTCGCTCAAGCGCACTCCCGTTCTCGATAATGTTTTTTCGCGCAGACATCAACCGTT	1380
Db	1321	TGTCGCTCAAGCGCACTCCCGTTCTCGATAATGTTTTTTCGCGCAGACATCAACCGTT	1380
Qy	1381	CTGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCTCGCTCGTATAATGTGGA	1440
Db	1381	CTGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCTCGCTCGTATAATGTGGA	1440
Qy	1441	ATTGTAGCGGATAACAATTTTACACAGGAAACAGCCAGTCCGTTTAGTGTTTTTACGA	1500
Db	1441	ATTGTAGCGGATAACAATTTTACACAGGAAACAGCCAGTCCGTTTAGTGTTTTTACGA	1500
Qy	1501	GCACTTCAACCAACGAAGCAACATAG-----	1524
Db	1501	GCACTTCAACCAACGAAGCAACATAGCAATGAAAAAATTAATTTCGCAATTCCTTTAGTG	1560
Qy	1525	-----ATTATGAAACTGAAGAGTAACTGTAATCTGG	1560
Db	1561	GTACTTTCTAATCTCACTCGGCCGATATGAAAAATCGAAGAGTAACCTGGTAATCTGG	1620
Qy	1561	ATTAAACCGCGATAAAGGCTATAACCGGCTCGCTGAAGTCGGTAAGAAATTCGAGAAGAT	1620
Db	1621	ATTAAACCGCGATAAAGGCTATAACCGGCTCGCTGAAGTCGGTAAGAAATTCGAGAAGAT	1680
Qy	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGAAGAGAAATTCACACAGTT	1680
Db	1681	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGAAGAGAAATTCACACAGTT	1740
Qy	1681	CGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTGGTGCGTAC	1740
Db	1741	CGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTGGTGCGTAC	1800
Qy	1741	GCTCAATCTGGCCTGTTGGCTGAAATCAACCCCGACAAAGCGTTCCAGGCAACAGCTGTAT	1800
Db	1801	GCTCAATCTGGCCTGTTGGCTGAAATCAACCCCGACAAAGCGTTCCAGGCAACAGCTGTAT	1860
Qy	1801	CCGTTTACCTGGATGCGGTACGTTACAAACGCAAGCTGATGTCCTTACCCGATCGCTGTT	1860
Db	1861	CCGTTTACCTGGATGCGGTACGTTACAAACGCAAGCTGATGTCCTTACCCGATCGCTGTT	1920
Qy	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTCCGACCCCGCAAAAACCTGGGAA	1920
Db	1921	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTCCGACCCCGCAAAAACCTGGGAA	1980
Qy	1921	GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGTAAAGCGCGCTGATGTTTCAAC	1980
Db	1981	GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGTAAAGCGCGCTGATGTTTCAAC	2040
Qy	1981	CTGCAAGAACCGTATTCTACCTGCGCGCTGATGCTCTGACGCGGGGTTATGCGTTCAAG	2040
Db	2041	CTGCAAGAACCGTATTCTACCTGCGCGCTGATGCTCTGACGCGGGGTTATGCGTTCAAG	2100

Qy	2041	TATGAAAACGGCAAGTACGACATTAAGAGCGTGGCGGTGGATTAACGCTGGCGGAAAGCG	2100
Db	2101	TATGAAAACGGCAAGTACGACATTAAGAGCGTGGCGGTGGATTAACGCTGGCGGAAAGCG	2160
Qy	2101	GGTCTGACCTTCCTCGTTGACCTGATTAAAAAACAACACATGAATGAGACACCGGATTAC	2160
Db	2161	GGTCTGACCTTCCTCGTTGACCTGATTAAAAAACAACACATGAATGAGACACCGGATTAC	2220
Qy	2161	TCCATCGCAGAAGCTGCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCGCCGTGG	2220
Db	2221	TCCATCGCAGAAGCTGCTTTAATAAAGGCGAAACAGCGATGACCATCAACGGCGCCGTGG	2280
Qy	2221	GCATGCTCCAACATCGACACAGCAAAAGTGAAATTAATGGTGTAAACGGTACTGCCGACCTTC	2280
Db	2281	GCATGCTCCAACATCGACACAGCAAAAGTGAAATTAATGGTGTAAACGGTACTGCCGACCTTC	2340
Qy	2281	AAGGGTCAACGATCCAAACCGTTGGTTGGCGTGTAGCGCAGGTATTAACGGCGCCAGT	2340
Db	2341	AAGGGTCAACGATCCAAACCGTTGGTTGGCGTGTAGCGCAGGTATTAACGGCGCCAGT	2400
Qy	2341	CCGAACAAAGAGCTGCGAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCG	2400
Db	2401	CCGAACAAAGAGCTGCGAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCG	2460
Qy	2401	GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG	2460
Db	2461	GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG	2520
Qy	2461	TTGCGGAAGATCCACGTATTGCGGCCACCATGGMAAACGCCAGAAAGTGAATCATG	2520
Db	2521	TTGCGGAAGATCCACGTATTGCGGCCACCATGGMAAACGCCAGAAAGTGAATCATG	2580
Qy	2521	CCGAACATCCCGCAGATGTCCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCC	2580
Db	2581	CCGAACATCCCGCAGATGTCCGCTTTCTGGTATGCCGTGCGTACTGCGGTGATCAACGCC	2640
Qy	2581	GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGTCG	2640
Db	2641	GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGTCG	2700
Qy	2641	AACAACAAACAATAACAATAACAACAACTCGGGATCGAGGGAAGGATTTTCAGAAATTC	2700
Db	2701	AACAACAAACAATAACAATAACAACAACTCGGGATCGAGGGAAGGATTTTCAGAAATTC	2760
Qy	2701	GGATCCTCT 2709	
Db	2761	GGATCCTCT 2769	

Search completed: August 1, 2005, 17:21:40
Job time : 9286 secs

THIS PAGE BLANK (USPTO)

Result No.	Score	Query %		DB	ID	Description
		Match	Length			
1	3300	100.0	3300	4	AD111591	Aad11591 Partial s
2	3252.8	98.0	3300	4	AD111592	Aad11592 Partial s
3	3215.2	97.4	3300	4	AD111588	Aad11588 Partial s
4	3215.2	97.4	3300	4	AD111589	Aad11589 Partial s
5	3207.2	97.2	3300	4	AD111590	Aad11590 Partial s
6	2705.8	82.0	8101	4	AAB25681	Aab25681 Nucleotid
7	2705.8	82.0	8101	6	AB153238	Ab153238 Nucleotid
8	2690.4	81.5	7259	12	ADO23608	Ado23608 DNA encod
9	2690.4	81.5	7322	12	ADO23613	Ado23613 DNA encod
10	2690.4	81.5	7352	12	ADO23603	Ado23603 DNA encod
11	2690.4	81.5	7370	12	ADO23598	Ado23598 DNA encod
12	2690.4	81.5	7370	12	ADO23639	Ado23639 DNA encod
13	2690.4	81.5	7370	12	ADO23649	Ado23649 DNA encod
14	2690.4	81.5	7370	12	ADO23644	Ado23644 DNA encod
15	2690.4	81.5	7403	12	ADO23594	Ado23594 DNA encod
16	2690.4	81.5	7442	12	ADO23590	Ado23590 DNA encod
17	2690.4	81.5	7478	12	ADO23584	Ado23584 DNA encod
18	2688.8	81.5	7112	12	ADO23618	Ado23618 DNA encod
19	2636.8	79.9	6806	6	AB149925	Ab149925 Maltose b
20	2605.4	79.0	7553	12	ADO23588	Ado23588 DNA encod

Db 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCGTGTGAGGTCGTAATCACTGTCATATTCG 1320
Qy 1321 TGTGCGTCAAGGCGACCTCCGTTCTGGATATATGTTTTTGGCGCGACATCATACGGTT 1380
Db 1321 TGTGCGTCAAGGCGACCTCCGTTCTGGATATATGTTTTTGGCGCGACATCATACGGTT 1380
Qy 1381 CTGGCAATATTTCTGAAATGAGCTGTGACAATTAATCATCGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAATATTTCTGAAATGAGCTGTGACAATTAATCATCGCTCGTATATGTTGGA 1440
Qy 1441 ATTGTGAGCGGATAACAATTTTACACAGAGAAACAGCCAGTCCGTTTGGTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAATTTTACACAGAGAAACAGCCAGTCCGTTTGGTTTTCACGA 1500
Qy 1501 GCACCTTCCAAACAGGACCATAGATTATGAAAACTGAAGAAGGTAAACTGTTATCTGG 1560
Db 1501 GCACCTTCCAAACAGGACCATAGATTATGAAAACTGAAGAAGGTAAACTGTTATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACGCTCTCGCTGAGTCCGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGCTCTCGCTGAGTCCGTAAGAAATTCAGAAAGAT 1620
Qy 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCACAGGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTTCTGGGCACACGACGCTTTGGTGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTTCTGGGCACACGACGCTTTGGTGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGTGGCTGAATCACCCCGGACAAAGCGTTCCAGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGTGGCTGAATCACCCCGGACAAAGCGTTCCAGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGATGGCGGTACGTTACAACGCGAAGCTGATTTGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGGCGGTACGTTACAACGCGAAGCTGATTTGCTTACCCGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATACAAAGATCTGCTGCGAAACCCGCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATACAAAGATCTGCTGCGAAACCCGCAAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGACTGAAGCGAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGACTGAAGCGAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCTGACTTACCTTGGCGCTGATTGCTGCTGACGGGGTTTATCGTTCAAG 2040
Db 1981 CTGCAAGAACCTGACTTACCTTGGCGCTGATTGCTGCTGACGGGGTTTATCGTTCAAG 2040
Qy 2041 TATGAAAACGGCAAGTACGACATTAAGACGTTGGCGTGATTAACGCTGGCGCAAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGTTGGCGTGATTAACGCTGGCGCAAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGTTGACCTGATTAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTGACCTGATTAAACAAACACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGAAGAGCTGCCCTTTAATAAGCGGAAACAGCGCATGACCATCAACCGCCGCTGG 2220
Db 2161 TCCATCGAAGAGCTGCCCTTTAATAAGCGGAAACAGCGCATGACCATCAACCGCCGCTGG 2220
Qy 2221 GCATGGTCCAAACATCGACACCGCAAGTGAATTTATGGTGAACGCTACTGCGACCTTC 2280
Db 2221 GCATGGTCCAAACATCGACACCGCAAGTGAATTTATGGTGAACGCTACTGCGACCTTC 2280
Qy 2281 AAGGTTCAACCATCCAAACCGTTTGGTTGGCTGCTGAGCGCAGGTATTAAACGCGCCAGT 2340
Db 2281 AAGGTTCAACCATCCAAACCGTTTGGTTGGCTGCTGAGCGCAGGTATTAAACGCGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGCGCAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGTCG 2400
Db 2341 CCGAACAAAGAGCTGCGCAAGAGTTCCTCGAAACTATCTGCTGACTGATGAAGGTCG 2400

Qy 2401 GAAGCGTTAATAAAGACAAACCGCTGGGTAGCGCTGAAGTCTTACGAGGAAGG 2460
Db 2401 GAAGCGTTAATAAAGACAAACCGCTGGGTAGCGCTGAAGTCTTACGAGGAAGG 2460
Qy 2461 TTGGGAAAAGATCCACGATTATGCGCCACCATGGAAGAACGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGGAAAAGATCCACGATTATGCGCCACCATGGAAGAACGCCAGAAAGGTGAAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCCGTCGTA CTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTTCTGGTATGCCGTCGTA CTGCGGTGATCAACGCC 2580
Qy 2581 GCCAGCGTGTGTCAGACTGTCGATGAAGCCCTGAAAGAACGCCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTGTGTCAGACTGTCGATGAAGCCCTGAAAGAACGCCAGACTAAATTCGAGCTCG 2640
Qy 2641 AACAAACAACAATAAACAACAACCACTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACAAACAACAATAAACAACAACCACTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Qy 2701 GGATCCTCTTCTCTGTGGCCAGCGGCCCTCGAGCCCCGGGAGAACCTTATGCTTGT 2760
Db 2701 GGATCCTCTTCTCTGTGGCCAGCGGCCCTCGAGCCCCGGGAGAACCTTATGCTTGT 2760
Qy 2761 CCGGAATGTGTAAGTCTTTCAGCCAGAGCAGCTCCCTGGTTCGCCACCCAGCGTACCCAC 2820
Db 2761 CCGGAATGTGTAAGTCTTTCAGCCAGAGCAGCTCCCTGGTTCGCCACCCAGCGTACCCAC 2820
Qy 2821 ACGGTTGAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGCCAGAGCAGCAGC 2880
Db 2821 ACGGTTGAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGCCAGAGCAGCAGC 2880
Qy 2881 CTGTTGCGCCATCAACGCACTCATCTGCGGAGAAAGCCATACAAATGTCCAGAAATGTGCG 2940
Db 2881 CTGTTGCGCCATCAACGCACTCATCTGCGGAGAAAGCCATACAAATGTCCAGAAATGTGCG 2940
Qy 2941 AAGTCTTTTCAAGTGTGCTGATCTTGGAGGACCAACAGTACTCAACCCGGGAGAAAG 3000
Db 2941 AAGTCTTTTCAAGTGTGCTGATCTTGGAGGACCAACAGTACTCAACCCGGGAGAAAG 3000
Qy 3001 CCCTATGCTTGTCCGGAATGTTAAGTCTTCTCTAGAGCTCTCACCCTGGTGGCCAC 3060
Db 3001 CCCTATGCTTGTCCGGAATGTTAAGTCTTCTCTAGAGCTCTCACCCTGGTGGCCAC 3060
Qy 3061 CAGCTTACCCACACCGGGTCAAAAACCGTATAAATGCCAGAGTGGCGGCAAAATCTTTTAGC 3120
Db 3061 CAGCTTACCCACACCGGGTCAAAAACCGTATAAATGCCAGAGTGGCGGCAAAATCTTTTAGC 3120
Qy 3121 CGCAGCATAACTGTTGGTGGCCATCAACGCACTCATCTGCGGAGAAAGCCATACAAATGT 3180
Db 3121 CGCAGCATAACTGTTGGTGGCCATCAACGCACTCATCTGCGGAGAAAGCCATACAAATGT 3180
Qy 3181 CCAGAAATGTCGCAAGTCTTCTCAACTTTCAGGCCATTTGGTCCGTCACCAAGTACTCAC 3240
Db 3181 CCAGAAATGTCGCAAGTCTTCTCAACTTTCAGGCCATTTGGTCCGTCACCAAGTACTCAC 3240
Qy 3241 ACCGTTAAAAAACTAGTGGCGCAGCGCCAGTAGTACCGCTAGCAGGTTCCGAGCTACGCT 3300
Db 3241 ACCGTTAAAAAACTAGTGGCGCAGCGCCAGTAGTACCGCTAGCAGGTTCCGAGCTACGCT 3300

RESULT 2

AAD11592

ID AAD11592 standard; DNA; 3300 BP.

XX AAD11592;

DT 24-SEP-2001 (first entry)

DE Partial sequence of pMal-Ap3 and ZFPap3 DNA.

KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;

KW modulation; plant technology; agriculture; Ap3; APETALA3; ds.
 OS Unidentified.

XX Key Location/Qualifiers
 XX CDS 2719..3270
 /tag= a
 /product= "ZFPap3 protein"
 /transl_except= (pos:3124..3126, aa:Ser)
 /transl_except= (pos:3127..3129, aa:Ser)
 /transl_except= (pos:3136..3138, aa:Val)
 /note= "CDS does not include start and stop codon"
 /partial
 /primer_bind 2740..2790
 /tag= b
 /bound_moiety= "F1-f2 primer"
 /bound_moiety= "F1-f1 primer"
 /tag= c
 /bound_moiety= "F2-b primer"
 /tag= d
 /bound_moiety= "F2-f primer"
 /tag= e
 /bound_moiety= "F3-b1 primer"
 /tag= f
 /bound_moiety= "F3-b2 primer"
 /tag= g
 /bound_moiety= "F4-f2 primer"
 /tag= h
 /bound_moiety= "F4-f1 primer"
 /tag= i
 /bound_moiety= "F5-b primer"
 /tag= j
 /bound_moiety= "F5-f primer"
 /tag= k
 /bound_moiety= "F6-b1 primer"
 /tag= l
 /bound_moiety= "F6-b2 primer"

WO200152620-A2.

26-JUL-2001.

19-JAN-2001; 2001WO-US001817.

21-JAN-2000; 2000US-0177468P.

21-JUL-2000; 2000US-00620897.

(SCRI) SCRIPPS RES INST.

(SYGN) SYNGENTA AGRIC DISCOVERY INC.

Barbas CF, Stege JT, Guan X, Dalmia B;

WPI; 2001-465325/50.

P-PSDB; AAE06005.

XX New zinc finger proteins, useful for modulating or regulating gene
 XX expression and metabolic pathways in plants, e.g. for treating in the
 XX plant cells a disorder that is associated with abnormal expression of the
 XX target gene.

XX Example 4; Page 148-149; 156pp; English.

XX The patent discloses methods and compositions to modulate the expression
 CC of a target gene in plant cells. The method involves providing plant
 CC cells with a zinc finger protein (ZFP) which is capable of specifically
 CC binding to a target nucleotide sequence or its complementary strand
 CC within a target gene and allowing the ZFP binding to the target
 CC nucleotide sequence, where the expression of the target gene in the plant
 CC cells is modulated. The ZFP and fusions of the ZFP proteins are useful
 CC for modulating or regulating gene expression and metabolic pathways in
 CC plants. The ZFP, fusion proteins and methods are useful in plant and
 CC agricultural technology. The method is useful particularly for treating a
 CC disorder in the plant cells, where the disorder is associated with
 CC abnormal expression of the target gene. The present DNA sequence is the
 CC partial sequence of pMal-Ap3 (APETALA3) and ZFPap3 DNA

XX Sequence 3300 BP; 853 A; 872 C; 876 G; 699 T; 0 U; 0 Other;

Query Match 98.0%; Score 3232.8; DB 4; Length 3300;
 Best Local Similarity 98.7%; Pred. No. 0;
 Matches 3258; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy	1	CCGACACCATCGAATGTGTGCAAAACCTTTCGCGTATGGCATGATAGCGCCCGGAAGAGA	60
Db	1	CCGACACCATCGAATGTGTGCAAAACCTTTCGCGTATGGCATGATAGCGCCCGGAAGAGA	60
Qy	61	GTCAATTCAAGGTGTGAATGTGAACCACTTAAGTATACGATGTCGAGAGTATCCCG	120
Db	61	GTCAATTCAAGGTGTGAATGTGAACCACTTAAGTATACGATGTCGAGAGTATCCCG	120
Qy	121	GTGTCCTTATCAGACCGTTTCCGCGTGTGAACCAAGCCAGCCAGCGTTCGCGAAAA	180
Db	121	GTGTCCTTATCAGACCGTTTCCGCGTGTGAACCAAGCCAGCCAGCGTTCGCGAAAA	180
Qy	181	CGCGGAAAGTGAAGCGCGATGCGGAGCTGAATACATTCGCGTGTGGCAGC	240
Db	181	CGCGGAAAGTGAAGCGCGATGCGGAGCTGAATACATTCGCGTGTGGCAGC	240
Qy	241	AACAATCGCGCGCAACAGTCGTGCTGATTGGCGTTGGCAGCTCCAGTCTGGCCCTGC	300
Db	241	AACAATCGCGCGCAACAGTCGTGCTGATTGGCGTTGGCAGCTCCAGTCTGGCCCTGC	300
Qy	301	ACGCGCGTGCAGAAATTTGTCGCGCGATTAATCTCGCGCGATCACTGGGTGCCAGCG	360
Db	301	ACGCGCGTGCAGAAATTTGTCGCGCGATTAATCTCGCGCGATCACTGGGTGCCAGCG	360
Qy	361	TGTTGTTGTCGATGTTAGAACGAGCGGCTCGAAGCGCTGAAAGCGGTGCAACAATC	420
Db	361	TGTTGTTGTCGATGTTAGAACGAGCGGCTGAAAGCGGTGCAACAATC	420
Qy	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACGAGATGCCA	480
Db	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACGAGATGCCA	480
Qy	481	TTGCTGTGGNAGCTGCTGCACTAATTTTCGCGCGTTATTTCTTGATGTCCTTGACGAGA	540
Db	481	TTGCTGTGGNAGCTGCTGCACTAATTTTCGCGCGTTATTTCTTGATGTCCTTGACGAGA	540
Qy	541	CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC	600
Db	541	CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATC	600
Qy	601	TGTTGCGATTGGGTCAACGAGCAAAATCGCGTGTAGCGGCGCCATTAAGTTCGTCTCGG	660
Db	601	TGTTGCGATTGGGTCAACGAGCAAAATCGCGTGTAGCGGCGCCATTAAGTTCGTCTCGG	660
Qy	661	CGGTCCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCCATAG	720
Db	661	CGGTCCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCCCATAG	720
Qy	721	CGGAACGGGAAGCGGACTGGAGTGCATGTCGCGTTTTCACAAACCAATGCAAAATGCTGA	780
Db	721	CGGAACGGGAAGCGGACTGGAGTGCATGTCGCGTTTTCACAAACCAATGCAAAATGCTGA	780


```
Db 2941 AAGTCTTTCAGCGAGTCCAGCAACCTGGTGGCCCAACCAACGTACTCACACCGGGGAGAG 3000
Qy 3001 CCCTATGCTTGTCCGGAATGTGGTAAGTCTTCTCTCAGAGCTCTCACCTGGTGGCCAC 3060
Db 3001 CCCTATGCTTGTCCGGAATGTGGTAAGTCTTCTCAGCAGCAGTGGCTCTTGGTTAGACAC 3060
Qy 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACCGGTGAAAAACCGTATAAATGCCAGAGTGGCGAAATCTTTTAGC 3120
Qy 3121 CGCAGCGATAACCTGGTGGCCATCAAGCGCACTCATCTAGTGGGAGAGCCATACAAATGT 3180
Db 3121 CAGCGCGCCACCTGGAGCGCCATCAAGCGCACTCATCTAGTGGGAGAGCCATACAAATGT 3180
Qy 3181 CCAGAATGTGGCAAGTCTTTCTCAACTTCAGGCCATTTGGTCCGTCACCAACGTACTCAC 3240
Db 3181 CCAGAATGTGGCAAGTCTTTCTCAACTTCAGGCCATTTGGTCCGTCACCAACGTACTCAC 3240
Qy 3241 ACCGGTAAAAAATAGTGGCGAGCGCGCCGCGCAGTACCGGTACGAGTTCCGGACTACGCT 3300
Db 3241 ACCGGTAAAAAATAGTGGCGAGCGCGCCGCGCAGTACCGGTACGAGTTCCGGACTACGCT 3300

RESULT 3
AAD11588
ID AAD11588 standard; DNA; 3300 BP.
XX
AC
XX AAD11588;
DT 24-SEP-2001 (first entry)
XX
DE Partial sequence of pMal-m1 and ZFPm1 DNA.
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
modulation; plant technology; agriculture; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT 2719..3270
FT CDS
FT /*tag= a
FT /product= "ZFPm1 protein"
FT /transl_except= (pos:2785..2787, aa:Asp)
FT /transl_except= (pos:2788..2790, aa:Pro)
FT /transl_except= (pos:2791..2793, aa:Gly)
FT /transl_except= (pos:2872..2874, aa:Arg)
FT /transl_except= (pos:2875..2877, aa:Ala)
FT /transl_except= (pos:2878..2880, aa:His)
FT /transl_except= (pos:2884..2886, aa:Glu)
FT /transl_except= (pos:2953..2955, aa:Gln)
FT /transl_except= (pos:2959..2961, aa:Ser)
FT /transl_except= (pos:3121..3123, aa:Arg)
FT /transl_except= (pos:3124..3126, aa:Ser)
FT /transl_except= (pos:3127..3129, aa:Asp)
FT /transl_except= (pos:3130..3132, aa:Asn)
FT /transl_except= (pos:3136..3138, aa:Val)
FT /transl_except= (pos:3139..3141, aa:Arg)
FT /transl_except= (pos:3205..3207, aa:Gln)
FT /transl_except= (pos:3208..3210, aa:Ala)
FT /transl_except= (pos:3211..3213, aa:Gly)
FT /transl_except= (pos:3214..3216, aa:His)
FT /transl_except= (pos:3220..3222, aa:Ala)
FT /transl_except= (pos:3223..3225, aa:Ser)
FT /note= "CDS does not include start and stop codon"
FT /partial
FT 2740..2790
FT /*tag= b
FT /bound_moiety= "F1-f2 primer"
FT 2770..2850
FT /*tag= c
FT /bound_moiety= "F1-f1 primer"
FT 2824..2889
FT /*tag= d
```

```
FT primer_bind
FT 2867..2940
FT /*tag= e
FT /bound_moiety= "F2-f primer"
FT 2916..2973
FT /*tag= f
FT /bound_moiety= "F3-b1 primer"
FT 2953..3021
FT /*tag= g
FT /bound_moiety= "F3-b2 primer"
FT 2992..3042
FT /*tag= h
FT /bound_moiety= "F4-f2 primer"
FT 3022..3102
FT /*tag= i
FT /bound_moiety= "F4-f1 primer"
FT 3076..3141
FT /*tag= j
FT /bound_moiety= "F5-b primer"
FT 3119..3192
FT /*tag= k
FT /bound_moiety= "F5-f primer"
FT 3168..3225
FT /*tag= l
FT /bound_moiety= "F6-b1 primer"
FT 3205..3273
FT /*tag= m
FT /bound_moiety= "F6-b2 primer"
XX
PN WO200152620-A2.
XX
XX 26-JUL-2001.
XX
XX 19-JAN-2001; 2001WO-US001817.
XX
XX 21-JAN-2000; 2000US-017468P.
XX 21-JUL-2000; 2000US-00620897.
XX
XX (SCRI ) SCRIPTS RES INST.
XX (SYGN ) SYNGENTA AGRIC DISCOVERY INC.
XX
XX Barbas CF, Stege JT, Guan X, Dalmia B;
XX WPI; 2001-465325/50.
XX P-PSDB; AAE06000.
XX
XX New zinc finger proteins, useful for modulating or regulating gene
XX expression and metabolic pathways in plants, e.g. for treating in the
XX plant cells a disorder that is associated with abnormal expression of the
XX target gene.
XX
XX Example 4; Page 138-139; 156pp; English.
XX
XX The patent discloses methods and compositions to modulate the expression
XX of a target gene in plant cells. The method involves providing plant
XX cells with a zinc finger protein (ZFP) which is capable of specifically
XX binding to a target nucleotide sequence or its complementary strand
XX within a target gene and allowing the ZFP binding to the target
XX nucleotide sequence, where the expression of the target gene in the plant
XX cells is modulated. The ZFP and fusions of the ZFP proteins are useful
XX for modulating or regulating gene expression and metabolic pathways in
XX plants. The ZFP, fusion proteins and methods are useful in plant and
XX agricultural technology. The method is useful particularly for treating a
XX disorder in the plant cells, where the disorder is associated with
XX abnormal expression of the target gene. The present DNA sequence is the
XX partial sequence of pMal-m1 and ZFPm1 DNA
XX
XX Sequence 3300 BP; 847 A; 874 C; 876 G; 703 T; 0 U; 0 Other;
SQ
```

```
Query Match 97.4%; Score 3215.2; DB 4; Length 3300;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 3247; Conservative 0; Mismatches 53; Indels 0; Gaps 0;
```

QY 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGGCATGATAGCCCGGAGAGA 60
Db 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGGCATGATAGCCCGGAGAGA 60
QY 61 GTCAATTCAAGGGTGGTGAATGTGA AAACCAAGTAAACGTTATACGATGTCGACAGTATGCCG 120
Db 61 GTCAATTCAAGGGTGGTGAATGTGA AAACCAAGTAAACGTTATACGATGTCGACAGTATGCCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCC CGCGTGGTGAACCAAGCGCAGCGCA CGTTTCTGCGAAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCC CGCGTGGTGAACCAAGCGCAGCGCA CGTTTCTGCGAAAA 180
QY 181 CGCGGAAAAAGTGAAGCGGCGATGG CGGAGCTGAATTACATTTCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGGCGATGG CGGAGCTGAATTACATTTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAAAACAGTCT GTTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300
Db 241 AACAACTGGCGGGCAAAACAGTCT GTTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGGCCCTGC 300
QY 301 ACGCGCGTCCGCAATTTGTCCG GCGATTAATCTCCGCGCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCGCGTCCGCAATTTGTCCG GCGATTAATCTCCGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTTCGATGGTAGAAGCAAG CGGCGTGAAGCGCTGTAAAGCGGCGGTGCAACAATC 420
Db 361 TGGTGGTTCGATGGTAGAAGCAAG CGGCGTGAAGCGCTGTAAAGCGGCGGTGCAACAATC 420
QY 421 TTCTCGCGCAACCGGTCA GTGGGTGATCAATTAACCTCCGTGGATCACAGGATGCCA 480
Db 421 TTCTCGCGCAACCGGTCA GTGGGTGATCAATTAACCTCCGTGGATCACAGGATGCCA 480
QY 481 TTGTCTGGAAAGTCCCTGC CTAATATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA 540
Db 481 TTGTCTGGAAAGTCCCTGC CTAATATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA 540
QY 541 CACCCATCAACAGTATTA TTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTA TTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
QY 601 TGGTCGATTTGGTCAC CAGAAATTCGCGCTGTTAGCGGGGCCATTAAGTTCTGTCTCGG 660
Db 601 TGGTCGATTTGGTCAC CAGAAATTCGCGCTGTTAGCGGGGCCATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTCGCTGTGGCTGG CATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTGTGGCTGG CATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGCGACTG GAGTGCCATGTCGGTTTTCAACAAACCATGCAATGCTGA 780
Db 721 CGGAACGGGAAGGCGACTG GAGTGCCATGTCGGTTTTCAACAAACCATGCAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCA CTGCGATGCTGGTTGCCAACGATCGATGGCGTGGGCGCAA 840
Db 781 ATGAGGGCATGTTCCCA CTGCGATGCTGGTTGCCAACGATCGATGGCGTGGGCGCAA 840
QY 841 TGGCGGCCATTAACAGAT CCGGGCTGCGCGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGGCCATTAACAGAT CCGGGCTGCGCGTTGGTGGCGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTC ATGTTATATCCCGCGTTTAAACCAACATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTC ATGTTATATCCCGCGTTTAAACCAACATCAAAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACG CTTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGGGTGA 1020
Db 961 GCCTGCTGGGCAACACG CTTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGTC CCGCTCTCACTGGTGAAAAAGAAACCAACCTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGTC CCGCTCTCACTGGTGAAAAAGAAACCAACCTGGCGGCCAATA 1080
QY 1081 CGCAAAACCGCCTCTCC CCGCGGCTTGGCGCGATTCAATTAATGCAGCTGGCAACGACAGGTTT 1140

Db 1081 CGCAAAACCGCCTCTCC CCGCGGCTTGGCGCGATTCAATTAATGCAGCTGGCAACGACAGGTTT 1140
QY 1141 CCCGACTGGAAAGCGGGC AGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAAGCGGGC AGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTGA CAGCTTATCATCGATGCAACGCTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGA CAGCTTATCATCGATGCAACGCTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGCGACCCATCGAAAG CTGTGATGCGTGTGCGTGTGAGTCTGAGTCTGATTAATCTG 1320
Db 1261 TCAGCGACCCATCGAAAG CTGTGATGCGTGTGCGTGTGAGTCTGAGTCTGATTAATCTG 1320
QY 1321 TGTGCGCTCAAGGCGCACT CCGGTTCTGGATAAATTTTTTTCGCGCGACATCAATACGGTT 1380
Db 1321 TGTGCGCTCAAGGCGCACT CCGGTTCTGGATAAATTTTTTTCGCGCGACATCAATACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAAT GAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTGCGA 1440
Db 1381 CTGGCAAAATTTCTGAAAT GAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTGCGA 1440
QY 1441 ATTCTGAGCGGATAACAAT TTTTTCACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
Db 1441 ATTCTGAGCGGATAACAAT TTTTTCACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCACGA 1500
QY 1501 GCACCTTCACCAACAAAGG ACCATAGATTATGAAAACTGAAGAAGGTAAACTGGTAACTCGG 1560
Db 1501 GCACCTTCACCAACAAAGG ACCATAGATTATGAAAACTGAAGAAGGTAAACTGGTAACTCGG 1560
QY 1561 ATTTAACGGCGATAAAGG CTATAACCGTCTCGCTGAAATCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTTAACGGCGATAAAGG CTATAACCGTCTCGCTGAAATCGGTAAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACG GTTGGACATCCCGATTAACCTGGAAGAGAAATTTCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACG GTTGGACATCCCGATTAACCTGGAAGAGAAATTTCCACAGGTT 1680
QY 1681 CGGCAACTGCGGATGGCC CTGACATTTCTTCTGGGCACACGACCGCTTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGCGGATGGCC CTGACATTTCTTCTGGGCACACGACCGCTTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTTG GCTGAAATCAACCCGGACAAAGCGTTCCAGGACAAGCTGAT 1800
Db 1741 GCTCAATCTGGCCTGTTG GCTGAAATCAACCCGGACAAAGCGTTCCAGGACAAGCTGAT 1800
QY 1801 CCGTTTACCTGGGATGCG GTACGTTTACAAACGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCG GTACGTTTACAAACGCAAGCTGATTGCTTACCCGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATT TTAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATT TTAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAA AGAAGCTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAA AGAAGCTGAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTT CACCTGCGCTGATGCTGCTGACCGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTT CACCTGCGCTGATGCTGCTGACCGGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACG ACATTAAGAGCTGGCGGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACG ACATTAAGAGCTGGCGGTGGATTAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGGTT GACCTGATTTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGGTT GACCTGATTTAAAAACAAACACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAAAGCTGC CTTTAAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220

Db 2161 TCCATCGCAGAGCTGCCCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCGGTGG 2220
Qy 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAATGTTAAACGGTACTGCCGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACAGCAAGTGAATTAATGTTAAACGGTACTGCCGACCTTC 2280
Qy 2281 AAGGTCAACCATCAAAACCGTTCGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTCAACCATCAAAACCGTTCGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Qy 2341 CGGAACAAGAGCTGGCAAAAGAGTCTCGAATAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CGGAACAAGAGCTGGCAAAAGAGTCTCGAATAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Qy 2461 TTGGCGAAAGATCCAGTATTGCGCCACCATGGAAGAAACGCCCAAGAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTGCGCCACCATGGAAGAAACGCCCAAGAGGTGAATCATG 2520
Qy 2521 CGGAACATCCGCGAGATGTCGCTTTCTGCTGATGCGGTGCTGCTGCTGATCAAGCC 2580
Db 2521 CGGAACATCCGCGAGATGTCGCTTTCTGCTGATGCGGTGCTGCTGCTGATCAAGCC 2580
Qy 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Qy 2641 AACACAACAATAAACAATAAACAACCTCGGGATCGAGGGAAGATTCAGAATTC 2700
Db 2641 AACACAACAATAAACAATAAACAACCTCGGGATCGAGGGAAGATTCAGAATTC 2700
Qy 2701 GGATCCTCTCTCTGTCGCGCAGCGCCCTCGAGCCCGGGAGAGCCCTATGCTGT 2760
Db 2701 GGATCCTCTCTCTGTCGCGCAGCGCCCTCGAGCCCGGGAGAGCCCTATGCTGT 2760
Qy 2761 CCGGAATGTGTAAGTCTTCTCTCAGAGCTCTCAGCTGTCGCCACAGCGTACCCAC 2820
Db 2761 CCGGAATGTGTAAGTCTTCTCTCAGAGCTCTCAGCTGTCGCCACAGCGTACCCAC 2820
Qy 2821 ACGGGTGAAAACCGTATAATGCCAGAGTGGCGGCAATCTTTAGCCAGAGCAGC 2880
Db 2821 ACGGGTGAAAACCGTATAATGCCAGAGTGGCGGCAATCTTTAGCCAGTCCAGCAAC 2880
Qy 2881 CTGGTGGCCATCAACGCACTCATACTGGCGAGAGCCATACAAATGTCCAGATGTGGC 2940
Db 2881 CTGGTGGCCATCAACGCACTCATACTGGCGAGAGCCATACAAATGTCCAGATGTGGC 2940
Qy 2941 AAGTCTTTCTCGGTCTGCAAAATCTGTCGCGCAGCAACCAAGTACTCACCGGGAGAAG 3000
Db 2941 AAGTCTTTCTCGGTCTGCAAAATCTGTCGCGCAGCAACCAAGTACTCACCGGGAGAAG 3000
Qy 3001 CCCTATGCTTGTCCGGAATGTGTAAGTCTTCTCAGAGCTCTCAGCTGTCGCCAC 3060
Db 3001 CCCTATGCTTGTCCGGAATGTGTAAGTCTTCTCAGAGCTCTCAGCTGTCGCCAC 3060
Qy 3061 CAGCGTACCCACACGGGTGAAAACCGTATAAATGCCAGAGTGGCGCAATCTTTTACG 3120
Db 3061 CAGCGTACCCACACGGGTGAAAACCGTATAAATGCCAGAGTGGCGCAATCTTTTACG 3120
Qy 3121 CGAGCGATACCTGGTGGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
Db 3121 CAGCCCGGCCATCTGGCGAGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
Qy 3181 CCAGATGTGGCAAGTCTTTCTCAACTTCAGGCCATTTGTCGCTACCAACGCTACTCAC 3240
Db 3181 CCAGATGTGGCAAGTCTTTCTCAACTTCAGGCCATTTGTCGCTACCAACGCTACTCAC 3240
Qy 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGGCAAGTACCCGTACGAGTTCGGGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCCAGCGCGGCAAGTACCCGTACGAGTTCGGGACTACGCT 3300

RESULT 4
AAD11589
ID AAD11589 standard; DNA; 3300 BP.
XX
AC AAD11589;
XX
DT 24-SEP-2001 (first entry)
XX
DE Partial sequence of pMal-m2 and ZFPm2 DNA.
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
XX
KW modulation; plant technology; agriculture; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
CDS 2719..3270
FT /*tag= a
FT /product= "ZFPm2 protein"
FT /note= "CDS does not include start and stop codon"
FT /partial
FT primer_bind 2740..2790
FT /*tag= b
FT /bound_moiety= "F1-f2 primer"
FT primer_bind 2770..2850
FT /*tag= c
FT /bound_moiety= "F1-f1 primer"
FT primer_bind 2824..2889
FT /*tag= d
FT /bound_moiety= "F2-b primer"
FT primer_bind 2867..2940
FT /*tag= e
FT /bound_moiety= "F2-f primer"
FT primer_bind 2916..2973
FT /*tag= f
FT /bound_moiety= "F3-b1 primer"
FT primer_bind 2953..3021
FT /*tag= g
FT /bound_moiety= "F3-b2 primer"
FT primer_bind 2992..3042
FT /*tag= h
FT /bound_moiety= "F4-f2 primer"
FT primer_bind 3022..3102
FT /*tag= i
FT /bound_moiety= "F4-f1 primer"
FT primer_bind 3076..3141
FT /*tag= j
FT /bound_moiety= "F5-b primer"
FT primer_bind 3119..3192
FT /*tag= k
FT /bound_moiety= "F5-f primer"
FT primer_bind 3168..3225
FT /*tag= l
FT /bound_moiety= "F6-b1 primer"
FT primer_bind 3205..3273
FT /*tag= m
FT /bound_moiety= "F6-b2 primer"

WO200152620-A2.

26-JUL-2001.

19-JAN-2001; 2001WO-US001817.

21-JAN-2000; 2000US-0177468P.

21-JUL-2000; 2000US-00620897.

(SCRI) SCRIPPS RES INST.

(SYGN) SYNGENTA AGRIC DISCOVERY INC.

Barbas CF, Stege JT, Guan X, Dalmia B;

XX WPI: 2001-465325/50.
DR P-PSDB; AAE06002.
XX
PT New zinc finger proteins, useful for modulating or regulating gene
PT expression and metabolic pathways in plants, e.g. for treating in the
PT plant cells a disorder that is associated with abnormal expression of the
PT target gene.
XX
PS Example 4; Page 140-142; 156pp; English.
XX
CC The patent discloses methods and compositions to modulate the expression
CC of a target gene in plant cells. The method involves providing plant
CC cells with a zinc finger protein (ZFP) which is capable of specifically
CC binding to a target nucleotide sequence or its complementary strand
CC within a target gene and allowing the ZFP binding to the target
CC nucleotide sequence, where the expression of the target gene in the plant
CC cells is modulated. The ZFP and fusions of the ZFP proteins are useful
CC for modulating or regulating gene expression and metabolic pathways in
CC plants. The ZFP, fusion proteins and methods are useful in plant and
CC agricultural technology. The method is useful particularly for treating a
CC disorder in the plant cells, where the disorder is associated with
CC abnormal expression of the target gene. The present DNA sequence is the
CC partial sequence of pMal-m2 and ZFPm2 DNA
XX
SQ Sequence 3300 BP; 847 A; 874 C; 876 G; 703 T; 0 U; 0 Other;

Query Match 97.4%; Score 3215.2; DB 4; Length 3300;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 3247; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 1 CCGACACCATCAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCCCGGAAGAGA 60
DB 1 CCGACACCATCAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCCCGGAAGAGA 60
QY 61 GTCATTCAGGGTGTGAATGTGAACACAGTAACGTATACGATGTCGAGAGTATGCG 120
DB 61 GTCATTCAGGGTGTGAATGTGAACACAGTAACGTATACGATGTCGAGAGTATGCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCAGCTTTCTCGAAAA 180
DB 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCCAGCCAGCTTTCTCGAAAA 180
QY 181 CCGGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTAATCAATCCCAACCGCGTGGCAC 240
DB 181 CCGGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTAATCAATCCCAACCGCGTGGCAC 240
QY 241 AACAACTGCGGGCAACAGTCTGTTGCTGATTTGGCGTTTGCACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGCGGGCAACAGTCTGTTGCTGATTTGGCGTTTGCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGTTCGCAAAATTTGTCGGGGCAATTAATCTCGCGCCGATCAACTGGGTGGCAGCG 360
DB 301 ACGCGCGTTCGCAAAATTTGTCGGGGCAATTAATCTCGCGCCGATCAACTGGGTGGCAGCG 360
QY 361 TGGTGGTGTGATGTAGAACGAAAGCGCGTGAAGCTGTAAAGCGCGGTGCAATC 420
DB 361 TGGTGGTGTGATGTAGAACGAAAGCGCGTGAAGCTGTAAAGCGCGGTGCAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATCTATCCGCTGGATGACAGGATGCCA 480
DB 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATCTATCCGCTGGATGACAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCATTAATGTTTCGGCGTTATTTCTGATGTTCTGACCCAGA 540
DB 481 TTGCTGTGGAAGCTGCTGCATTAATGTTTCGGCGTTATTTCTGATGTTCTGACCCAGA 540
QY 541 CACCCATCAACAGTATTTATTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTTATTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
QY 601 TGGTGGCATGGGTACCGAAGAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
DB 601 TGGTGGCATGGGTACCGAAGAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660

DB 601 TGGTGGCATGGGTACCGAAGAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTGCGTCTGCGTGGCATAAATATCTCATCTCGCAATCAAAATTCAGCGCATAG 720
DB 661 CGCGTCTGCGTCTGCGTGGCATAAATATCTCATCTCGCAATCAAAATTCAGCGCATAG 720
QY 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCCGGTTTTCAACAACCAATGCAAAATGCTGA 780
DB 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCCGGTTTTCAACAACCAATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGTCACACGATCAGATGCGCTGGCGCGAA 840
DB 781 ATGAGGGCATCGTTCCCACTCGCATGCTGTTGTCACACGATCAGATGCGCTGGCGCGAA 840
QY 841 TGGCGGCCATTAACGAGTCCCGGCTGCGCGTGGTGGCGATATCTCGGTAGTGGGATAG 900
DB 841 TGGCGGCCATTAACGAGTCCCGGCTGCGCGTGGTGGCGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCAATGCAAAATGCTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCAATGCAAAATGCTTC 960
QY 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
DB 961 GCCTGCTGGGCAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCGTCTCATCTGTTGAAAGAAAAACCACTGGCGCCCAATA 1080
DB 1021 AGGGCAATCAGCTGTTGCCCGTCTCATCTGTTGAAAGAAAAACCACTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGATTAATTAATGAGTGTAGTCTACTCATTTAG 1140
DB 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGATTAATTAATGAGTGTAGTCTACTCATTTAG 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTGTAGTCTACTCATTTAG 1200
DB 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTGTAGTCTACTCATTTAG 1200
QY 1201 GCACAAATTCATGTTTGAAGCTTATCATCGACTGCGAGTGCACCAATGTTCTGGCG 1260
DB 1201 GCACAAATTCATGTTTGAAGCTTATCATCGACTGCGAGTGCACCAATGTTCTGGCG 1260
QY 1261 TCAGGCAGCATCGGAAGCTGTGATGCTGTGCGAGTGTAAATCACTGCTAATTTTCG 1320
DB 1261 TCAGGCAGCATCGGAAGCTGTGATGCTGTGCGAGTGTAAATCACTGCTAATTTTCG 1320
QY 1321 TGTGCTCAAGCGGCGACTCCCGTTCTGGAATAATGTTTTTGGCGCGACATATAACGGTT 1380
DB 1321 TGTGCTCAAGCGGCGACTCCCGTTCTGGAATAATGTTTTTGGCGCGACATATAACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGAACAATTAATCATCGGCTCGTATAATGTTGGA 1440
DB 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGAACAATTAATCATCGGCTCGTATAATGTTGGA 1440
QY 1441 ATTGTAGCGGATTAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTTTTCACGA 1500
DB 1441 ATTGTAGCGGATTAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGGTTTTCACGA 1500
QY 1501 GCACCTTACCAACAGGACCATAGATTATGAAAACCTGAAGAGGTAATACTGGTAATCTGG 1560
DB 1501 GCACCTTACCAACAGGACCATAGATTATGAAAACCTGAAGAGGTAATACTGGTAATCTGG 1560
QY 1561 ATTAACGGCGATAAAGGCTATAACCGGCTCGCTGAAGTCCGTAAGAAATTCAGAGAAAGAT 1620
DB 1561 ATTAACGGCGATAAAGGCTATAACCGGCTCGCTGAAGTCCGTAAGAAATTCAGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTACCGGTTGAGCATCCGGAATAAATCTGGAAGAGAAATTCAGAGGTT 1680
DB 1621 ACCGGAATTAAGTACCGGTTGAGCATCCGGAATAAATCTGGAAGAGAAATTCAGAGGTT 1680
QY 1681 CGGGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTGTTGGCTAC 1740
DB 1681 CGGGCAACTGCGGATGGCCCTGACATTAATCTTCTGGGCACACGCGCTTTGTTGGCTAC 1740


```
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCCCGGCAAAAGCGTTTCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCCCGGCAAAAGCGTTTCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCGCGTACGTTTAAACGGCAAGCTGATTCACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGATGCGCGTACGTTTAAACGGCAAGCTGATTCACCCGATCGCTGT 1860
QY 1861 GAAGCGTTATCGCTGATTTATAAACAAGATCTGCTGCGCAACCCGCCAAAACTTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAAACAAGATCTGCTGCGCAACCCGCCAAAACTTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGAGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGAGGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCCTACTTCACTGGCGCGTGTATGCTGTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCCTACTTCACTGGCGCGTGTATGCTGTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGCGCTGGATAAGCGTGGCGGCAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGCGCTGGATAAGCGTGGCGGCAAGCG 2100
QY 2101 GGTCTGAOCTTCTCGTTGACCTGATTAAGAAACAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGAOCTTCTCGTTGACCTGATTAAGAAACAAACACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTAAATAAGCGCAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTAAATAAGCGCAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCAACATCGACACACAGCAAAAGTGAATATGGTGAACCGTACTCCGACCTTC 2280
Db 2221 GCATGGTCAACATCGACACACAGCAAAAGTGAATATGGTGAACCGTACTCCGACCTTC 2280
QY 2281 AAGGTCAACCATCAAAACCGTTGCTGGCGTGTAGCGCAGGTATTAACGCCCGCAGT 2340
Db 2281 AAGGTCAACCATCAAAACCGTTGCTGGCGTGTAGCGCAGGTATTAACGCCCGCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGGTTCTCCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGGTTCTCCGAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGCTTTACGAGGAAG 2460
Db 2401 GAAGCGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGCTTTACGAGGAAG 2460
QY 2461 TTGGGAAAGATCCACGTATTGCGCGCACCATGGAAGAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGGAAAGATCCACGTATTGCGCGCACCATGGAAGAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTCTGCTGATGTCGCTGCGTACTCGCGTGATCAAGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTCTGCTGATGTCGCTGCGTACTCGCGTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACACACACAAATACAAATACAACTCCGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACACACACAAATACAAATACAACTCCGGATCGAGGGAAGGATTCAGAAATTC 2700
QY 2701 GGATCCTCTTCTGTCGCGCAGCGCGCTTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCCTCTTCTGTCGCGCAGCGCGCTTCGAGCCCGGGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTGTAAAGTCTTCAGCGCAGACAGCTCCTGCTGCGCCACACGCGTACCAC 2820
Db 2761 CCGGAATGTGTAAAGTCTTCCTCTCAGAGCTCTCACCTGCTGCGCCACACGCGTACCAC 2820
```

RESULT 5

```
AAD11590
ID AAD11590 standard; DNA; 3300 BP.
XX
AC AAD11590;
DT 24-SEP-2001 (first entry)
XX
DE Partial sequence of pMal-m3 and ZFPm3 DNA.
XX
KW Zinc finger protein; ZFP; gene expression; metabolic pathway regulator;
  modulation; plant technology; agriculture; ds.
XX
OS Unidentified.
XX
FH Key
FT CDS
FT 2719..3270
FT /*tag= a
FT /product= "ZFPm3 protein"
FT /note= "CDS does not include start and stop codon"
FT /partial
FT /*tag= b
FT /bound_moiety= "F1-f2 primer"
FT 2770..2850
FT /*tag= c
FT /bound_moiety= "F1-f1 primer"
FT 2824..2889
FT /*tag= d
FT /bound_moiety= "F2-b primer"
FT 2867..2940
FT /*tag= e
FT /bound_moiety= "F2-f primer"
FT 2916..2973
FT /*tag= f
FT /bound_moiety= "F3-b1 primer"
FT 2953..3021
FT /*tag= g
FT /bound_moiety= "F3-b2 primer"
FT 2992..3042
FT primer_bind
```

FT		/*tag= h	
FT	primer_bind	/bound_moiety= "F4-f2 primer"	
FT		3022. .3102	
FT		/*tag= i	
FT	primer_bind	/bound_moiety= "F4-f1 primer"	
FT		3076. .3141	
FT		/*tag= j	
FT	primer_bind	/bound_moiety= "F5-b primer"	
FT		3119. .3192	
FT		/*tag= k	
FT	primer_bind	/bound_moiety= "F5-f primer"	
FT		3168. .3225	
FT		/*tag= l	
FT	primer_bind	/bound_moiety= "F6-b1 primer"	
FT		3205. .3273	
FT		/*tag= m	
FT	primer_bind	/bound_moiety= "F6-b2 primer"	
XX			
XX	WO200152620-A2.		
XX			
XX	26-JUL-2001.		
XX			
XX	19-JAN-2001; 2001WO-US001817.		
XX			
XX	21-JAN-2000; 2000US-0177468P.		
PR	21-JUL-2000; 2000US-00620897.		
XX			
XX	(SCRI) SCRIPPS RES INST.		
PA	(SYGN) SYNGENTA AGRIC DISCOVERY INC.		
PA			
XX			
PI	Barbas CF, Stege JT, Guan X, Dalmia B;		
PI			
XX	WPI; 2001-465325/50.		
DR	P-PSDB; AAE06003.		
DR			
XX			
XX	New zinc finger proteins, useful for modulating or regulating gene		
PT	expression and metabolic pathways in plants, e.g. for treating in the		
PT	plant cells a disorder that is associated with abnormal expression of the		
PT	target gene.		
XX			
XX	Example 4; Page 143-145; 156pp; English.		
XX			
CC	The patent discloses methods and compositions to modulate the expression		
CC	of a target gene in plant cells. The method involves providing plant		
CC	cells with a zinc finger protein (ZFP) which is capable of specifically		
CC	binding to a target nucleotide sequence or its complementary strand		
CC	within a target gene and allowing the ZFP binding to the target		
CC	nucleotide sequence, where the expression of the target gene in the plant		
CC	cells is modulated. The ZFP and fusions of the ZFP proteins are useful		
CC	for modulating or regulating gene expression and metabolic pathways in		
CC	plants. The ZFP, fusion proteins and methods are useful in plant and		
CC	agricultural technology. The method is useful particularly for treating a		
CC	disorder in the plant cells, where the disorder is associated with		
CC	abnormal expression of the target gene. The present DNA sequence is the		
CC	Partial sequence of pMal-m3 and ZFPm3 DNA		
XX			
XX	Sequence 3300 BP; 845 A; 878 C; 877 G; 700 T; 0 U; 0 Other;		
XX			
XX	Query Match	97.2%; Score 3207.2; DB 4; Length 3300;	
XX	Best Local Similarity	98.2%; Pred. No. 0;	
XX	Matches 3242; Conservative	0; Mismatches 58; Indels 0; Gaps 0;	
Qy	1	CCGACACCATCGAATGGTGTGCAAAACCTTTCCGGGTATGCGCATGATGAGCCCGGAAGACA	60
Db	1	CCGACACCATCGAATGGTGTGCAAAACCTTTCCGGGTATGCGCATGATGAGCCCGGAAGACA	60
Qy	61	GTCATTCAGGGTGTGTAATGTGTGAACACAGTAACGTTATACGATGTCGCAGAGTATGCCG	120
Db	61	GTCATTCAGGGTGTGTAATGTGTGAACACAGTAACGTTATACGATGTCGCAGAGTATGCCG	120
Qy	121	GTGTCCTTATCAGACCGTTTCCCGGTGGTGAAACAGGCCAGCCACGTTTCTGGCAAAA	180
Db	121	GTGTCCTTATCAGACCGTTTCCCGGTGGTGAAACAGGCCAGCCACGTTTCTGGCAAAA	180

```
QY 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGCAGGTGTAATCACTGCATTAATTCG 1320
DB 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGCAGGTGTAATCACTGCATTAATTCG 1320
QY 1321 TGTCCGCTCAAGGGGCACCTCCGCTTCTGGATATGTTTTTGGCCGCAATCATACGGTT 1380
DB 1321 TGTCCGCTCAAGGGGCACCTCCGCTTCTGGATATGTTTTTGGCCGCAATCATACGGTT 1380
QY 1381 CTGGCAAAATATTCTGAAATGAGCTGTGACAATTAATCATCGCTCGTATATGTTGGA 1440
DB 1381 CTGGCAAAATATTCTGAAATGAGCTGTGACAATTAATCATCGCTCGTATATGTTGGA 1440
QY 1441 ATTGTAGCGGATACAATTTCAACAGAAACAGCCAGTCGGTTAGGTGTTTTACGA 1500
DB 1441 ATTGTAGCGGATACAATTTCAACAGAAACAGCCAGTCGGTTAGGTGTTTTACGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAATGAAAATCTGAAGAGGTAATCGTAACTCTGG 1560
DB 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAATGAAAATCTGAAGAGGTAATCGTAACTCTGG 1560
QY 1561 ATTAAACGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTAAACGCGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCTCAAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAATTCCTCAAGGTT 1680
QY 1681 GCGGCACTGCGGATGCGCCCTGACATTATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
DB 1681 GCGGCACTGCGGATGCGCCCTGACATTATCTTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGGCTGAATCAACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
DB 1741 GCTCAATCTGGCTGTGGCTGAATCAACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCGGTACGTTTAAACGGCAAGCTGATGCTTACCCGATCGCTGTT 1860
DB 1801 CCGTTTACCTGGATGCGGTACGTTTAAACGGCAAGCTGATGCTTACCCGATCGCTGTT 1860
QY 1861 GAAGGTTATCGCTGATTATTAACAAAGATCTGCTGCCGACCCCGCAAAACCTCGGAA 1920
DB 1861 GAAGGTTATCGCTGATTATTAACAAAGATCTGCTGCCGACCCCGCAAAACCTCGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATTAAGAACTGAACCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
DB 1921 GAGATCCCGGCGCTGGATTAAGAACTGAACCGAAAGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
DB 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGGCAAGTACGACATTAAAGACGTGGCGTGGATACGCTGGCGGAAAGCG 2100
DB 2041 TATGAAAAACGGCAAGTACGACATTAAAGACGTGGCGTGGATACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAAACAAACAAACATGAATGAGACACCGATTAC 2160
DB 2101 GGTCTGACCTTCTCGTTGACCTGATTAAACAAACAAACATGAATGAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAAGAGCGAACACAGGATGACCATCAACGGCCCGTGG 2220
DB 2161 TCCATCGCAGAGCTGCTTTTAAAGAGCGAACACAGGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCCAACATCGACACCGCAAAAGTGAATTAATGTTGTAACCGTACCTGCCACCTTC 2280
DB 2221 GCATGGTCCCAACATCGACACCGCAAAAGTGAATTAATGTTGTAACCGTACCTGCCACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTGCGTGGCGTGTGAGCGAGGATTAATACGCGCCAGT 2340
DB 2281 AAGGGTCAACCATCCAAACCGTTGCGTGGCGTGTGAGCGAGGATTAATACGCGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTTCGAAAACTATCTGCTGACTGATGAAGGCTG 2400
```

```
DB 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGTGACTGATGAAGGCTG 2400
QY 2401 GAAAGGTTTAATAAAGACAAACCCGCTGGGTGCGTAGCGCTGAAAGTCTTACAGGAAAG 2460
DB 2401 GAAAGGTTTAATAAAGACAAACCCGCTGGGTGCGTAGCGCTGAAAGTCTTACAGGAAAG 2460
QY 2461 TTGGGGAAGATCCACGTTATTTGCCGCCACCATGAAAAACGCCAGAAAGGTGAATCATG 2520
DB 2461 TTGGGGAAGATCCACGTTATTTGCCGCCACCATGAAAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGAGTGTCCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
DB 2521 CCGAACATCCCGCAGAGTGTCCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTGTGATGATGATGAAGCCCTGAAACGCGCAGACTAAATTCGAGCTG 2640
DB 2581 GCCAGCGTGTGATGATGATGAAGCCCTGAAACGCGCAGACTAAATTCGAGCTG 2640
QY 2641 AACAAACAAACAAATAACAACAACTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
DB 2641 AACAAACAAACAAATAACAACAACTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
QY 2701 GGATCTCTTCTCTGTGGCCAGCGCCCTCGAGCCCGGGGAGAAAGCCCTATGCTTGT 2760
DB 2701 GGATCTCTTCTCTGTGGCCAGCGCCCTCGAGCCCGGGGAGAAAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTGTAAGTCTCTTACGACGAGAGGCTCTCTGTGGGCCACACGAGCGTACCAC 2820
DB 2761 CCGGAATGTGTAAGTCTCTTACGACGATCTCTGGCCACCTGGTTGCCACGAGCTACCAC 2820
QY 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGCCAGAGCAGCAG 2880
DB 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTTAGCCAGAGCAGCAG 2880
QY 2881 CTGGTGGCCATCAACGCACTCATCTGCGGAGAGGCCATACAAATGTCAGAAATGTGGC 2940
DB 2881 CTGGTGGCCATCAACGCACTCATCTGCGGAGAGGCCATACAAATGTCAGAAATGTGGC 2940
QY 2941 AAGTCTTTAGTGTGTCGTGATCTTTCAGAGGACCAACGTACTCACACGGGGGAGAAG 3000
DB 2941 AAGTCTTTAGTGTGTCGTGATCTTTCAGAGGACCAACGTACTCACACGGGGGAGAAG 3000
QY 3001 CCTATGCTTGTCCGGAATGTGTAAGTCTCTCTCAGAGCTCTCACCTGTGGCGCCAC 3060
DB 3001 CCTATGCTTGTCCGGAATGTGTAAGTCTTTCAGCAGAGCAGCTCCCTGGTGGCCAC 3060
QY 3061 CAGGTAACCCACACCGGTGAAAAACCGTATAAATGCCAGAGTCCGCAAAATCTTTTAGT 3120
DB 3061 CAGGTAACCCACACCGGTGAAAAACCGTATAAATGCCAGAGTCCGCAAAATCTTTTAGT 3120
QY 3121 CGCAGGATTAACCTGTGGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
DB 3121 GACTGCGGACCTTGTCTGCCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
QY 3181 CCAGAAATGTGGCAAGTCTTTCTCAACTTCAGGCGATTTGGTCCGTCACCAACGTAAC 3240
DB 3181 CCAGAAATGTGGCAAGTCTTTCTCCCAATCCAGGCACTCTGTCGGCAGCCACCACTAC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCCAGCGCCAGTACCCGTAACCGTTCGCGACTACGCT 3300
DB 3241 ACCGGTAAAAAACTAGTGGCCAGCGCCAGTACCCGTAACCGTTCGCGACTACGCT 3300
```

RESULT 6

AAH25681

ID AAH25681 standard; DNA; 8101 BP.

XX

AC AAH25681;

XX

DT 05-SEP-2001 (first entry)

XX

DE Nucleotide sequence of human Cdc25C phosphatase DNA.

XX Fusion protein; maltose binding protein; MBP; Cdc25C phosphatase;
 KW Cdc25B1; Cdc25B2; Cdc25B3; cyclin-dependent kinase; CDK; cell division;
 KW ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200144467-A2.
 XX
 PD 21-JUN-2001.
 XX
 XX 13-DEC-2000; 2000WO-FR003496.
 XX
 PF 14-DEC-1999; 99FR-00015722.
 PR 30-MAY-2000; 2000FR-00006883.
 PR 21-SEP-2000; 2000FR-00012008.
 XX
 XX (SCRC) SCRAS SOC CONSEILS RECH & APPL SCI.
 XX
 XX Goubin-Gramatica F, Ducommun B, Prevost G;
 XX WPI; 2001-398152/42.
 XX
 PT New fusion protein of maltose-binding protein and Cdc phosphatase, useful
 PT for identifying phosphatase modulators for regulating the cell cycle.
 XX
 PS Claim 7; Page 9-13; 56pp; French.
 XX
 CC The present sequence encodes human Cdc25C phosphatase, and was expressed
 CC in Escherichia coli. It is used to make fusion proteins with Escherichia
 CC coli maltose binding protein (MBP). The specification describes fusion
 CC proteins of MBP and one of Cdc 25B1, 25B2, 25B3 and 25C. Cdc phosphatases
 CC are involved in activation of cyclin-dependent kinases (CDK) that control
 CC cell division. The fusion proteins are used to identify modulators of the
 CC specified human Cdc phosphatases, potentially useful for regulation of
 CC cell division. They are also for studying physiological or
 CC physiopathological activities of Cdc phosphatase
 XX
 SQ Sequence 8101 BP; 2127 A; 1987 C; 2121 G; 1866 T; 0 U; 0 Other;
 Query Match 82.0%; Score 2705.8; DB 4; Length 8101;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 CGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGCGATGATAGCCCGGGAAGA 60
 DB 1 CGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGCGATGATAGCCCGGGAAGA 60
 QY 61 GTC AATTCAGGGTGTGTAATGTGAACCAACGTAACGTTATACGATGTCGAGAGTATGCCG 120
 DB 61 GTC AATTCAGGGTGTGTAATGTGAACCAACGTAACGTTATACGATGTCGAGAGTATGCCG 120
 QY 121 GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACCGACCGCACCGCTTTCTGCGA AAA 180
 DB 121 GTGTCCTTATCAGACCGTTTCCCGGTGGTGGTAACCGACCGCACCGCTTTCTGCGA AAA 180
 QY 181 CGCGGGA AAAAGTGAAGCGCGATGGCGGAGCTGAATTTACATTC CCAACCGCGTGGCAC 240
 DB 181 CGCGGGA AAAAGTGAAGCGCGATGGCGGAGCTGAATTTACATTC CCAACCGCGTGGCAC 240
 QY 241 AACAACTGGCGGCAAAACAGCTGTTGCTGATGGGGTGGCCACCTCGAGCTGGCCCTGC 300
 DB 241 AACAACTGGCGGCAAAACAGCTGTTGCTGATTTGGGGTGGCCACCTCGAGCTGGCCCTGC 300
 QY 301 ACGCCCGTTCGCA AATTTGTCGGGGGATTAATCTCGCGCGCATCAACTGGGTGCCAGCG 360
 DB 301 ACGCCCGTTCGCA AATTTGTCGGGGGATTAATCTCGCGCGCATCAACTGGGTGCCAGCG 360
 QY 361 TGGTGGTGTGCAATGTAGAAACGAAGCGCGTGC AAGCCCTGTAAAGCGGCGGTGCACAATC 420
 DB 361 TGGTGGTGTGCAATGTAGAAACGAAGCGCGTGC AAGCCCTGTAAAGCGGCGGTGCACAATC 420
 QY 421 TTCTCGCGAAACGCGTCAGTGGGCTGATCACTAACTATCCCGTGGATGACCAAGGATGCCA 480

Db
 421 TTCTCGCGCAACGGCTAGTGGGCTGATCACTAATATCCGCTGGATGACCAAGATGCCA 480
 QY
 481 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGCTCTGTGACCA 540
 DB
 481 TTGCTGTGGAAGCTGCCCTGCACATAATGTTCCGGCGTTATTTCTTGATGCTCTGTGACCA 540
 QY
 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
 DB
 541 CACCCATCAACAGTATTTATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
 QY
 601 TGGTGGCATTTGGGCTACCGCAAAATCGGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
 DB
 601 TGGTGGCATTTGGGCTACCGCAAAATCGGCTGTTAGCGGGCCCATTAAGTTCTGTCTCGG 660
 QY
 661 CGGCTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
 DB
 661 CGGCTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
 QY
 721 CGGAACGGGAAGGCGACTGGAGTGGCATGTCGGTTTTCAACAAACCATGCAATGCTGA 780
 DB
 721 CGGAACGGGAAGGCGACTGGAGTGGCATGTCGGTTTTCAACAAACCATGCAATGCTGA 780
 QY
 781 ATGAGGCGATGTTCCCACTCGCATGCTGTTGCCAAGATGCGCTGGGCGCA 840
 DB
 781 ATGAGGCGATGTTCCCACTCGCATGCTGTTGCCAAGATGCGCTGGGCGCA 840
 QY
 841 TSGCGCCATTCACGAGTCCGGGCTGGCGTGGTGGCGGATATCTCGTAGTGGGATAG 900
 DB
 841 TSGCGCCATTCACGAGTCCGGGCTGGCGTGGTGGCGGATATCTCGTAGTGGGATAG 900
 QY
 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACACCATCAACAGGATTTTC 960
 DB
 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACACCATCAACAGGATTTTC 960
 QY
 961 GCCTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCGACGGGTGA 1020
 DB
 961 GCCTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCGACGGGTGA 1020
 QY
 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGTGTAAGAAAAAACCCCTGGCGCCCAATA 1080
 DB
 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGTGTAAGAAAAAACCCCTGGCGCCCAATA 1080
 QY
 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAACAGAGTTT 1140
 DB
 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAACAGAGTTT 1140
 QY
 1141 CCGGACTGGAAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTTAGTCACTCATTTAG 1200
 DB
 1141 CCGGACTGGAAAGCGGCGAGTGAGCGCAACCAATTAATGTAGTTAGTCACTCATTTAG 1200
 QY
 1201 GCACAAATTCATGTTTGACAGCTTATCATCGACTGCAACGGGTGACCAATGTTCTGGCG 1260
 DB
 1201 GCACAAATTCATGTTTGACAGCTTATCATCGACTGCAACGGGTGACCAATGTTCTGGCG 1260
 QY
 1261 TCAGGACGCCATTCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAATCACTGCAATTTTCG 1320
 DB
 1261 TCAGGACGCCATTCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAATCACTGCAATTTTCG 1320
 QY
 1321 TGTGCTCAAGCGGCGACTCCCGTTCCTGGATAATGTTTTTGGCGCGACATCAACGGTT 1380
 DB
 1321 TGTGCTCAAGCGGCGACTCCCGTTCCTGGATAATGTTTTTGGCGCGACATCAACGGTT 1380
 QY
 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGCACAATTAATCATCGGCTCGTATAATGTGTGA 1440
 DB
 1381 CTGGCAAAATTTCTGAAATGAGCTGTTGCACAATTAATCATCGGCTCGTATAATGTGTGA 1440
 QY
 1441 ATTGTAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCAGA 1500
 DB
 1441 ATTGTAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCAGA 1500
 QY
 1501 GCACCTTACCACAAAGGACCATAGTTATGAAAACCTGAAGAGGTAAACTGGTATCTCG 1560

Db 1501 GCATTCACCAACAAAGGACCATAGATTATGAAAAATCGAAGAGGTAAATCTGGTATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTTATACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTTATACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTACCGTTCAGCATCCGATCGGATCGGATCGGATCGGATCGGATCGGATCGGAT 1680
Db 1621 ACCGGAATTAAGTACCGTTCAGCATCCGATCGGATCGGATCGGATCGGATCGGATCGGATCGGAT 1680
Qy 1681 GCGCAATCGCGATCGCGCTGACATTAATCTCTGGGACACAGCCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAATCGCGATCGCGCTGACATTAATCTCTGGGACACAGCCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGGCTGCTGGCTGAATCAACCGGACAAAGGCTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGGCTGCTGGCTGAATCAACCGGACAAAGGCTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTTACACCGCAAGCTGATTCCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACACCGCAAGCTGATTCCTTACCCGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATAAAGAGATCTGCTGCGAACCCGCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATAAAGAGATCTGCTGCGAACCCGCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTCATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTCATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATTCGCTGACGGGGGTTATGCTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATTCGCTGACGGGGGTTATGCTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTAGTAGCATTAAGACCTGGCGTGGATACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTAGTAGCATTAAGACCTGGCGTGGATACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGGTGTGCTGATTAATAAACAACACACATGAATGACACACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTGTGCTGATTAATAAACAACACACATGAATGACACACCGGATTAC 2160
Qy 2161 TCCATCCGAGAGCTGCTTTAATAAGGCGAACAACAGGATGACCATCAACGGCCGCTGG 2220
Db 2161 TCCATCCGAGAGCTGCTTTAATAAGGCGAACAACAGGATGACCATCAACGGCCGCTGG 2220
Qy 2221 GCATGGTCCAACTCGACACAGCAAGTGAATTAATGCTGTAACGGTACTGCGACCTTC 2280
Db 2221 GCATGGTCCAACTCGACACAGCAAGTGAATTAATGCTGTAACGGTACTGCGACCTTC 2280
Qy 2281 AAGGGTCAACATCCAAACCGTTTGGTGGCTGCTGAGCGCAGGATTAACGGCCGCGAGT 2340
Db 2281 AAGGGTCAACATCCAAACCGTTTGGTGGCTGCTGAGCGCAGGATTAACGGCCGCGAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAGAGTCTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAGAGTCTCTCGAAACTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGCTGCTGAGCGCAGGATTAACGGCCGAGT 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGTGGCTGCTGAGCGCAGGATTAACGGCCGAGT 2460
Qy 2461 TTGGCGAAGATCCACGTTATGCGCCACCATGGAACCGCCAGAAAGTGGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGTTATGCGCCACCATGGAACCGCCAGAAAGTGGAATCATG 2520
Qy 2521 CCGAACATCCCGAGATGCTCGCTTTCTGCTATGCGCTGGTACTGCGGATCAACGCC 2580
Db 2521 CCGAACATCCCGAGATGCTCGCTTTCTGCTATGCGCTGGTACTGCGGATCAACGCC 2580
Qy 2581 GCGAGCGGTGCTGAGTCTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGTCTG 2640
Db 2581 GCGAGCGGTGCTGAGTCTGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGTCTG 2640

Qy 2641 AACCAACAACAATTAACAACAACCTCGGGATCGAGGAAGGATTTCAAGATTC 2700
Db 2641 AACCAACAACAATTAACAACAACCTCGGGATCGAGGAAGGATTTCAAGATTC 2700
Qy 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709

RESULT 7
ABLS3238

ID ABL53238 standard; DNA; 8101 BP.

XX ABL53238;

XX 17-JUN-2002 (first entry)

XX Nucleotide sequence of pMAL-Hs Cdc25C.

XX Cdc25C phosphatase; amine; Cdc25 phosphatase; spontaneous alopecia;
XX alopecia; proliferative disease; parasitic disease; viral infection;
XX neurodegenerative disease; myopathy; angiogenesis; psoriasis; restenosis;
XX vitamin K; ss.

OS Synthetic.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1528..2677

FT /*tag= b

FT /note= "encodes maltose binding protein (MBP)"

FT 2713..4134

FT /*tag= a

FT /note= "ORF of human Cdc25C"

XX FR2812198-A1.

XX 01-FEB-2002.

XX 28-JUL-2000; 2000FR-00009900.

XX 28-JUL-2000; 2000FR-00009900.

XX (SCRC) SCRAS SOC CONSEILS RECH & APPL SCI.

XX Prevost G, Brezak PMC, Galcera CMO, Thuriereau C, Goubin GF;
XX Ducommun B;

XX WPI; 2002-166066/22.

XX cdc25 Phosphatase inhibiting medicaments, useful e.g. for treating tumor
XX diseases, viral infections, neurodegenerative disease or alopecia,
XX containing new or known aryl-substituted amines.

XX Disclosure; Page 18-22; 39pp; French.

XX The present sequence represents pMAL-HsCdc25C, a plasmid encoding human
XX Cdc25C phosphatase. The specification describes the use of aryl-
XX substituted secondary or tertiary amines for the production of
XX medicaments for inhibiting Cdc25 phosphatases, especially Cdc25C
XX phosphatases. The amines of the invention are used for treating
XX spontaneous alopecia or alopecia induced by exogenous products or
XX radiation. They are also used for treating tumoral or non-tumoral
XX proliferative diseases, parasitic diseases, viral infections,
XX neurodegenerative diseases or myopathy. Examples of non-tumoral
XX proliferative diseases are e.g. angiogenesis, psoriasis or restenosis.
XX The amines are also useful for inhibiting the proliferation of
XX microorganisms (especially yeasts) and in the treatment of all disorders
XX conventionally treated with vitamin K or its derivatives

XX Sequence 8101 BP; 2127 A; 1987 C; 2121 G; 1866 T; 0 U; 0 Other;

Query Match									
Best Local Similarity 82.0%; Score 2705.8; DB 6; Length 8101;									
Matches 2707; Conservative 0; Mismatches 2; Indels 0; Gaps 0;									
QY	1	CCGACACCATGAAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA	60						
Db	1	CCGACACCATGAAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA	60						
QY	61	GTCAATTACAGGGTGTGAATGTGAACACAGTAAACGTTATACGATCTCGAGAGTATGCG	120						
Db	61	GTCAATTACAGGGTGTGAATGTGAACACAGTAAACGTTATACGATCTCGAGAGTATGCG	120						
QY	121	GTGTCCTCTATCAGACCGTTTCCCGGTGGTGAACAGCCAGCCAGCCAGCTTTCTCGAATA	180						
Db	121	GTGTCCTCTATCAGACCGTTTCCCGGTGGTGAACAGCCAGCCAGCCAGCTTTCTCGAATA	180						
QY	181	CGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTAACATCCCAACCGCTGGGCAC	240						
Db	181	CGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTAACATCCCAACCGCTGGGCAC	240						
QY	241	AACAACTCGCGGCAAAAACAGTGTGCTGATTTGGCGTTGCCACTTCCAGTCTGGCCCTGC	300						
Db	241	AACAACTCGCGGCAAAAACAGTGTGCTGATTTGGCGTTGCCACTTCCAGTCTGGCCCTGC	300						
QY	301	ACGCGCGCTCGCAANTTGTCCGGCGATTAATCTCGCGCGATCAACTGGGTGCCAGG	360						
Db	301	ACGCGCGCTCGCAANTTGTCCGGCGATTAATCTCGCGCGATCAACTGGGTGCCAGG	360						
QY	361	TGGTGTGTCGATGTAGAAACGAGCGCGTGAAGCCTGTAAAGCGCGGTGCACAATC	420						
Db	361	TGGTGTGTCGATGTAGAAACGAGCGCGTGAAGCCTGTAAAGCGCGGTGCACAATC	420						
QY	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGTGGATGACCAAGGATGCCA	480						
Db	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGTGGATGACCAAGGATGCCA	480						
QY	481	TTGCTGTGGAGCTGCCCTGCACTAATGTTCGGCGTTATTTCTTGTATGTCTGTGACCA	540						
Db	481	TTGCTGTGGAGCTGCCCTGCACTAATGTTCGGCGTTATTTCTTGTATGTCTGTGACCA	540						
QY	541	CACCCATCAACAGTATTTTCTCCATGAAGACGGTACGGACTGGGCGTGGAGCATC	600						
Db	541	CACCCATCAACAGTATTTTCTCCATGAAGACGGTACGGACTGGGCGTGGAGCATC	600						
QY	601	TGGTCGCATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCG	660						
Db	601	TGGTCGCATTTGGGTCAACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCG	660						
QY	661	CGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
Db	661	CGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
QY	721	CGGAACGGAGGCGACATGGAGTGCATCTCCGGTTTTCACAAACCATGCAAAATGCTGA	780						
Db	721	CGGAACGGAGGCGACATGGAGTGCATCTCCGGTTTTCACAAACCATGCAAAATGCTGA	780						
QY	781	ATGAGGGCATCTTCCCACTGCGATGCTGGTGGCCACGATCAGATGGCGTGGCGCA	840						
Db	781	ATGAGGGCATCTTCCCACTGCGATGCTGGTGGCCACGATCAGATGGCGTGGCGCA	840						
QY	841	TGCGGCGCATTTACAGATCCCGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATACG	900						
Db	841	TGCGGCGCATTTACAGATCCCGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATACG	900						
QY	901	ACGATACCGAAGACAGCTCATGTATATCCCGCGCTTAACACCATCAAAACAGGATTTTC	960						
Db	901	ACGATACCGAAGACAGCTCATGTATATCCCGCGCTTAACACCATCAAAACAGGATTTTC	960						
QY	961	GCCTGCTGGGGCAACACAGCGTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA	1020						
Db	961	GCCTGCTGGGGCAACACAGCGTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA	1020						
QY	1021	AGGCAATACAGCTGTGGCCCGTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA	1080						

Db	1021	AGGCAATACAGCTGTGGCCCGTCTCACTGGTGAAGAAAAACCAACCTGGCGCCCAATA	1080						
QY	1081	CGCAAAACCGCTCTCCCGCGCGTGTGGCCGATTCAATTAATGACGTGGCAGCACAGGTTT	1140						
Db	1081	CGCAAAACCGCTCTCCCGCGCGTGTGGCCGATTCAATTAATGACGTGGCAGCACAGGTTT	1140						
QY	1141	CCGACGTGGAAGCGGGCAGTGAAGGCAACGCAATTAATGTAGTTAGTACTACTCATTTAG	1200						
Db	1141	CCGACGTGGAAGCGGGCAGTGAAGGCAACGCAATTAATGTAGTTAGTACTACTCATTTAG	1200						
QY	1201	GCACAAATTCATGTGTTGACAGCTTATCATGACGTGACAGGTGACCAATGCTTCTGGCG	1260						
Db	1201	GCACAAATTCATGTGTTGACAGCTTATCATGACGTGACAGGTGACCAATGCTTCTGGCG	1260						
QY	1261	TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGTAATCACTGTCATTAATTCG	1320						
Db	1261	TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGTAATCACTGTCATTAATTCG	1320						
QY	1321	TGTGCTCAAGCGGCACTCCCGTTCCTGGATTAATGTTTTTTCGCGCGGACATCAATACGGTT	1380						
Db	1321	TGTGCTCAAGCGGCACTCCCGTTCCTGGATTAATGTTTTTTCGCGCGGACATCAATACGGTT	1380						
QY	1381	CTGGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGTGGA	1440						
Db	1381	CTGGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTGTGGA	1440						
QY	1441	ATTGTGACGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTTCGCGCGGACATCAATACGGTT	1500						
Db	1441	ATTGTGACGGATTAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTTCGCGCGGACATCAATACGGTT	1500						
QY	1501	GCATTTCACCAAGGACCATAGATTATGAAAACTGAAGAAAGTAAACTGTGTAATTCGG	1560						
Db	1501	GCATTTCACCAAGGACCATAGATTATGAAAACTGAAGAAAGTAAACTGTGTAATTCGG	1560						
QY	1561	ATTAAACGGGATAAAGGCTAAACGGTCTCCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT	1620						
Db	1561	ATTAAACGGGATAAAGGCTAAACGGTCTCCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT	1620						
QY	1621	ACCGGAATTTAAAGTACACCGTTTTCGAGCATCCGGATTAATTCGGAAGAGAAATTCGCCAGGTT	1680						
Db	1621	ACCGGAATTTAAAGTACACCGTTTTCGAGCATCCGGATTAATTCGGAAGAGAAATTCGCCAGGTT	1680						
QY	1681	CGGCAACTGGCGATGGCCCTTGACATTAATCTTCGGGACACGACCGCTTTCGTTGGTGGCTAC	1740						
Db	1681	CGGCAACTGGCGATGGCCCTTGACATTAATCTTCGGGACACGACCGCTTTCGTTGGTGGCTAC	1740						
QY	1741	GCTCAATCTGGCGCTTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGTAT	1800						
Db	1741	GCTCAATCTGGCGCTTGGCTGAAATCAACCGGACAAAGCGTTCCAGGACAAAGCTGTAT	1800						
QY	1801	CCGTTTACCTGGGATGGCGTACGTTTACAAACGGCAAGCTGATTGCTTACCGATCGCTGTT	1860						
Db	1801	CCGTTTACCTGGGATGGCGTACGTTTACAAACGGCAAGCTGATTGCTTACCGATCGCTGTT	1860						
QY	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCGCCAAAAACCTGGGAA	1920						
Db	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCGCCAAAAACCTGGGAA	1920						
QY	1921	GAGATCCCGCGCTGGATTAAGAACTGAAGGAAAGGTGAAGGCGCGCTGATGTTCAAC	1980						
Db	1921	GAGATCCCGCGCTGGATTAAGAACTGAAGGAAAGGTGAAGGCGCGCTGATGTTCAAC	1980						
QY	1981	CTGCAAGAAACCGTACTTCCACTGGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG	2040						
Db	1981	CTGCAAGAAACCGTACTTCCACTGGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG	2040						
QY	2041	TATGAAAAACGCAAGTACGACATTAAGACAGTGGGCGTGGATTAACGCTGGCGCGAAACG	2100						
Db	2041	TATGAAAAACGCAAGTACGACATTAAGACAGTGGGCGTGGATTAACGCTGGCGCGAAACG	2100						
QY	2101	GGTCTGACCTTCTGGTTGACCTGATTTAAAAACAAACATGAATGAGACACCCGATTAC	2160						


```
Db 2101 GGTCTGACCTTCTCGTGGTTGACCTGATTTAAATAAACAACATGAATGAACACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCCTTTTAAATAAGCGGAAACAGCGATGACCATCAACGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCCTTTTAAATAAGCGGAAACAGCGATGACCATCAACGCGCCGCTGG 2220
QY 2221 GCATGGTCCACATCGACACACAGCAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCACATCGACACACAGCAAGTGAATTAATGTTGTAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTCTGCGCTGCTGAGCGAGGTATTAAGCGCCCGCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTCTGCGCTGCTGAGCGAGGTATTAAGCGCCCGCAGT 2340
QY 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGGAAAGATCCACGATTGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGGAAAGATCCACGATTGCGCCACCATGGAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTGATGCGGTGCTACTGCGGTGATCAAGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTGATGCGGTGCTACTGCGGTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGTCTG 2640
Db 2581 GCCAGCGTCTGACACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGTCTG 2640
QY 2641 AACCAACAACAATAACAATAACAACAACCTCGGGATCGAGGGAAGGATTTTCAAGATTC 2700
Db 2641 AACCAACAACAATAACAATAACAACAACCTCGGGATCGAGGGAAGGATTTTCAAGATTC 2700
QY 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709

RESULT 8
AD023608
ID AD023608 standard; DNA; 7259 BP.
AC
AC AD023608;
XX
XX 01-JUL-2004 (first entry)
XX
XX DNA encoding MBP-Toxop30del18(82-294aa) fusion protein.
XX
XX P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; Gene.
XX
XX Toxoplasma gondii.
OS Synthetic.
XX
XX US2004067239-A1.
XX
XX 08-APR-2004.
XX
XX 02-OCT-2002; 2002US-00263153.
XX
XX 02-OCT-2002; 2002US-00263153.
XX
XX (MAIN/) MAINE G T.
PA (FATE/) FATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX
```

PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;

XX WPI; 2004-304563/28.

DR P-FSDB; ADO23609.

XX

PT Novel purified polypeptide having sequence identity to amino acid

PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,

PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of Igm or Igg

PT antibodies to Toxoplasma gondii.

XX Example 2; Fig 22; 114pp; English.

PS

XX The invention relates to a purified P30 antigen (I) chosen from 3 fully

CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6

CC amino acids added to the C-terminus of the amino acid sequence of

CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in

CC which at least one of the five C-terminal cysteine amino acids of the

CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted

CC with alanine, or comprising the amino acid sequence chosen from MBP-

CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1

CC fusion proteins. (I) is useful for detecting the presence of Igm

CC antibodies to Toxoplasma gondii in a test sample, which involves

CC contacting the test sample suspected of containing the Igm antibodies

CC with a composition comprising (I) and detecting the presence of (I)/Igm

CC antibody complexes. The present sequence represents DNA encoding a MBP-

CC ToxoP30 fusion protein of the invention.

XX

SQ Sequence 7259 BP; 1870 A; 1822 C; 1926 G; 1641 T; 0 U; 0 Other;

Query Match

Best Local Similarity 81.5%; Score 2690.4; DB 12; Length 7259;

Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGGCATGATAGCGCCCGAAGAGA 60

Db 1 CCACACCATCGAATGGTGCAGAAACCTTTTCGGGTATGGCATGATAGCGCCCGAAGAGA 60

QY 61 GTCAAATTCAGGCTGGTGAATGTGAAACCAAGTATACGATGTCGAGAGTATGCCG 120

Db 61 GTCAAATTCAGGCTGGTGAATGTGAAACCAAGTATACGATGTCGAGAGTATGCCG 120

QY 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCGCCAGCCAGCTTTCTGCGAAA 180

Db 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACACGAGCGCCAGCCAGCTTTCTGCGAAA 180

QY 181 CGCGGGAAGAGTGAAGCGCGATGCGGAGCTGAATACATTCACCAACCGCGTGGCAC 240

Db 181 CGCGGGAAGAGTGAAGCGCGATGCGGAGCTGAATACATTCACCAACCGCGTGGCAC 240

QY 241 AACAACTGCGCGGCAACACAGTCTGCTGATTTGGCGTTGGCCACTCCAGTCTGCGCCTGC 300

Db 241 AACAACTGCGCGGCAACACAGTCTGCTGATTTGGCGTTGGCCACTCCAGTCTGCGCCTGC 300

QY 301 ACGCCGCTGCGMAATGTCGCGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 360

Db 301 ACGCCGCTGCGMAATGTCGCGCGGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG 360

QY 361 TGGTGTGTGATGTTAGAACGCGGTGCGAGCTGTAAGCGCGGTGTAAGCGCGGTGCAAAATC 420

Db 361 TGGTGTGTGATGTTAGAACGCGGTGCGAGCTGTAAGCGCGGTGTAAGCGCGGTGCAAAATC 420

QY 421 TTCTCGCGCAACCGCTGAGTGGGCTGATCAATTAATCTATCCGCTGGATGACAGGATGCCA 480

Db 421 TTCTCGCGCAACCGCTGAGTGGGCTGATCAATTAATCTATCCGCTGGATGACAGGATGCCA 480

QY 481 TTGCTGTGAAGCTGCTGCACTAATGTTCCGCGGTATTTCTTGATGTCTGACACAGA 540

Db 481 TTGCTGTGAAGCTGCTGCACTAATGTTCCGCGGTATTTCTTGATGTCTGACACAGA 540

QY 541 CACCCATCAACAGTATTTTCTCCCATGAAGAGGTACGCGACTGGGCGTGGAGCATC 600

Db 541 CACCCATCAACAGTATTTTCTCCCATGAAGAGGTACGCGACTGGGCGTGGAGCATC 600

QY 601 TGGTCGCAATTCGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGCTCGG 660
Db 601 TGGTCGCAATTCGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGCTCGG 660
QY 661 CGCGTCTCGGCTCGGCTGGGATGAATAATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db 661 CGCGTCTCGGCTCGGCTGGGATGAATAATCTCACTCGCAATCAAAATTCAGCGGATAG 720
QY 721 CGGAAACGGGAAGGCGATCGGAGTGCATGTCCGFTTTTCAACCAATGCAAAATGCTGA 780
Db 721 CGGAAACGGGAAGGCGATCGGAGTGCATGTCCGFTTTTCAACCAATGCAAAATGCTGA 780
QY 781 ATGAGGGCATGTTCCCACTCGGATGCTGTTGCCAAGATCAGATGCGCGTGGCGCAA 840
Db 781 ATGAGGGCATGTTCCCACTCGGATGCTGTTGCCAAGATCAGATGCGCGTGGCGCAA 840
QY 841 TGGCGGCATTAACCGAGTCCGGGCTGCGGTTGGTGCGGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGGCATTAACCGAGTCCGGGCTGCGGTTGGTGCGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACCGAAAGACAGCTCATGTTATATCCGCGGTTTAAACCAATCAAAACGAGATTTC 960
Db 901 ACGATACCGAAAGACAGCTCATGTTATATCCGCGGTTTAAACCAATCAAAACGAGATTTC 960
QY 961 GCCTGCTGGGCAACACAGCGTGGACCGGTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACAGCGTGGACCGGTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGTCGCGCTCTCACTGGTGAAAGAAACACACCTGGCGGCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGTCGCGCTCTCACTGGTGAAAGAAACACACCTGGCGGCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGATTCATTAATGCAGCTGGCAGCAGAGTTT 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGATTCATTAATGCAGCTGGCAGCAGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGTAGTACTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACCGCAATTAATGTAGTACTCACTCATTTAG 1200
QY 1201 GCACAATTCATGTTTGACAGCTTATCATGACCTGCACGCTGCACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTCATGTTTGACAGCTTATCATGACCTGCACGCTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGCAGCCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTAATCACTGCATAATTTCG 1320
Db 1261 TCAGCAGCCCATCGGAAGCTGTGGTATGGCTGTGCAGGTCGTAATCACTGCATAATTTCG 1320
QY 1321 TGTGCTCAAGCGCAGCTCCCGTTCTGGATAATGTTTTTGGCGCGCATCATCAACGGTT 1380
Db 1321 TGTGCTCAAGCGCAGCTCCCGTTCTGGATAATGTTTTTGGCGCGCATCATCAACGGTT 1380
QY 1381 CTGGCAAAATTTCTGAAATGAGCTGTGCAATTAATCATCGGCTCGTAAATGTTGTGA 1440
Db 1381 CTGGCAAAATTTCTGAAATGAGCTGTGCAATTAATCATCGGCTCGTAAATGTTGTGA 1440
QY 1441 ATTGTGAGGGATAACAATTTCAACAGAGAACAGCCAGTCGTTTAGGTGTTTCAAGA 1500
Db 1441 ATTGTGAGGGATAACAATTTCAACAGAGAACAGCCAGTCGTTTAGGTGTTTCAAGA 1500
QY 1501 GCACCTTCAACAAAGGACCATAGATTATGAAAACTGAAGAAAGGTAATCTGGTAACTCGG 1560
Db 1501 GCACCTTCAACAAAGGACCATAGCATATGAAAACTGAAGAAAGGTAATCTGGTAACTCGG 1560
QY 1561 ATTAACGGGATAAAGGCTATAACGCTCGCTGAGTTCGTTAGGTAATTCAGAAAGT 1620
Db 1561 ATTAACGGGATAAAGGCTATAACGCTCGCTGAGTTCGTTAGGTAATTCAGAAAGT 1620
QY 1621 ACCGGAATTAAGTCAACGCTGAGCATCCGATAAACTGGAAGAGAAATTCACAGAGTT 1680
Db 1621 ACCGGAATTAAGTCAACGCTGAGCATCCGATAAACTGGAAGAGAAATTCACAGAGTT 1680
QY 1681 GCGGCAACTGGCGATGGGCCCTGACATTATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740

RESULT 9
ADO23613
ID ADO23613 standard; DNA; 7322 BP.
XX AC ADO23613;

Db 1581 GCGGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACAGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGSCCTGTGTTGGCTGAAATCACCCCGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGSCCTGTGTTGGCTGAAATCACCCCGACAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCCGTACGTTACAAACGGCAAGCTGATGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCCGTACGTTACAAACGGCAAGCTGATGCTTACCCGATCGCTGTT 1860
QY 1861 GAACGCTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
Db 1861 GAACGCTTATCGCTGATTTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAGAACTGAAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATTTGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGCTGGCGTGGATAACGCTGGCGCGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAAACAAACACATGATGACAGACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAAACAAACACATGATGACAGACCGGATTAC 2160
QY 2161 TCCATCGCAAGAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCGCTGG 2220
Db 2161 TCCATCGCAAGAGCTGCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCGCTGG 2220
QY 2221 GCATGCTCAACATCGACACAGCAAAAGTGAATTTATGCTGTAACGCTGCTGCGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACAGCAAAAGTGAATTTATGCTGTAACGCTGCTGCGACCTTC 2280
QY 2281 AAGGCTCAACCATCCAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGCTCAACCATCCAAACCGTTCTGTTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAACATATCTGCTGACGTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAACATATCTGCTGACGTGATGAAGTCTG 2400
QY 2401 GAACGGTTAATAAGACAAACCGCTGGGTCCGCTAGCGCTGAAAGTCTTACGAGGAAG 2460
Db 2401 GAACGGTTAATAAGACAAACCGCTGGGTCCGCTAGCGCTGAAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGTTATTTGCCGCCACCATGGAACAAACGCCCAAGAGGTGAAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATTTGCCGCCACCATGGAACAAACGCCCAAGAGGTGAAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCGTTTCTGGTATGTCGCGTGTGATCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCGTTTCTGGTATGTCGCGTGTGATCTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACACTGTCGATGAAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACACTGTCGATGAAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAAATAACAAACAAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACAAACAAACAAATAACAAACAAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700

XX
DT 01-JUL-2004 (first entry)
XX
DE DNA encoding MBP-Toxop30del110(52-284aa) fusion protein.
XX
KW P30 antigen; Toxo30del13C; Toxo30del12C; Toxo30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; gene.
XX
OS Toxoplasma gondii.
OS Synthetic.
XX
PN US2004067239-A1.
XX
PD 08-APR-2004.
XX
PF 02-OCT-2002; 2002US-00263153.
XX
PR 02-OCT-2002; 2002US-00263153.
XX
PA (MAIN/) MAINE G T.
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX
PI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
XX
DR WPI; 2004-304563/28.
DR P-P8DB; ABO23614.
XX
PT Novel purified polypeptide having sequence identity to amino acid
PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
PT Toxo30del12C, Toxo30 MIX1, useful for detecting presence of IgM or IgG
PT antibodies to Toxoplasma gondii.
XX
PS Example 2; Fig 25; 114pp; English.
XX
CC The invention relates to a purified p30 antigen (I) chosen from 3 fully
CC defined Toxo30del13C, Toxo30del12C and Toxo30 MIX1 sequences, having 1-6
CC amino acids added to the C-terminus of the amino acid sequence of
CC Toxo30del12C p30 antigen sequence, or comprising an amino acid sequence in
CC which at least one of the five C-terminal cysteine amino acids of the
CC amino acid sequence of Toxo30del13C p30 antigen sequence is substituted
CC with alanine, or comprising the amino acid sequence chosen from MBP-
CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
CC fusion proteins. (I) is useful for detecting the presence of IgM
CC antibodies to Toxoplasma gondii in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC Toxo30 fusion protein of the invention.
SQ Sequence 7322 BP; 1887 A; 1844 C; 1934 G; 1657 T; 0 U; 0 Other;

Query Match 81.5%; Score 2690.4; DB 12; Length 7322;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGGCATGATAGCCCGCGGAAGAGA 60
DB 1 CCGACACCATCGAATGGTGC AAAACCTTTCCGGGTATGGCATGATAGCCCGCGGAAGAGA 60
QY 61 GTCAATTCAGGTGGTGAATGTGAAACAGTAACGTTATACGATGTCGAGAGATGCCG 120
DB 61 GTCAATTCAGGTGGTGAATGTGAAACAGTAACGTTATACGATGTCGAGAGATGCCG 120
QY 121 GTGTCTCTTATCAGACCGCTTTCCCGCGTGTGAACACGAGCCAGCCAGCTTTCTGCCGAAA 180
DB 121 GTGTCTCTTATCAGACCGCTTTCCCGCGTGTGAACACGAGCCAGCCAGCTTTCTGCCGAAA 180
QY 181 CGCGGGAAAAAGTGGAAAGCGCGGAGTGGCGAGCTGAATTACATTCCTCCAAACCGCGTGGCAC 240
DB 181 CGCGGGAAAAAGTGGAAAGCGCGGAGTGGCGAGCTGAATTACATTCCTCCAAACCGCGTGGCAC 240

QY 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGCAAAACAGTCGTTGCTGATTGGCGTTGCCACTCCAGTCTGGCCCTGC 300
QY 301 ACGCGCGCTCGCAAAATTTGTTCGCGCGGATTAATCTCCGCGCGGATCAACTGGGTCCGACG 360
DB 301 ACGCGCGCTCGCAAAATTTGTTCGCGCGGATTAATCTCCGCGCGGATCAACTGGGTCCGACG 360
QY 361 TGGTGGTGTGATGGTAGAAACGAAGCGCGTFCGAAGCCTGTAAAGCGCGGTGCACAATC 420
DB 361 TGGTGGTGTGATGGTAGAAACGAAGCGCGTFCGAAGCCTGTAAAGCGCGGTGCACAATC 420
QY 421 TTCTTCGCGCAACCGCTCAGTGGGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
DB 421 TTCTTCGCGCAACCGCTCAGTGGGCTGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
QY 481 TTGCTGTGGAAGCTGCCCTGCACATAATTTCCGGCGTTTATTTCTTGATGTCTCTGACCAGA 540
DB 481 TTGCTGTGGAAGCTGCCCTGCACATAATTTCCGGCGTTTATTTCTTGATGTCTCTGACCAGA 540
QY 541 CACCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
DB 541 CACCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC 600
QY 601 TGGTCGCATTTGGGTACCCAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGCTCGG 660
DB 601 TGGTCGCATTTGGGTACCCAGCAAAATCGCGCTGTAGCGGGCCCAATTAAGTTCTGCTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAAACGGGAAGCGGACTGAGTGCATGTCGCGTTTCAACAAACCATCAAAATGCTGA 780
DB 721 CGGAAACGGGAAGCGGACTGAGTGCATGTCGCGTTTCAACAAACCATCAAAATGCTGA 780
QY 781 ATGAGGCAATCGTTTCCCACTGCGATGCTGTTGCCAACGATCAGATGGCGCTGGCGCAA 840
DB 781 ATGAGGCAATCGTTTCCCACTGCGATGCTGTTGCCAACGATCAGATGGCGCTGGCGCAA 840
QY 841 TGGCGCGCATTAACGAGTCCGGCTCGGCTGGTTCGGTATCTCGGTAGTGGGATACG 900
DB 841 TGGCGCGCATTAACGAGTCCGGCTCGGCTGGTTCGGTATCTCGGTAGTGGGATACG 900
QY 901 ACCATACCGAAGACAGCTCATGTTATATCCGCGGTTAACCCACCATCAACAGGATTTTC 960
DB 901 ACCATACCGAAGACAGCTCATGTTATATCCGCGGTTAACCCACCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGGCAAAACAGCGTGCAGCCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
DB 961 GCCTGCTGGGGCAAAACAGCGTGCAGCCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCGCGTCTCACTGGTGAAGAAAACCAACCCCTGGCGCCCAATA 1080
DB 1021 AGGGCAATCAGCTGTTGCGCGTCTCACTGGTGAAGAAAACCAACCCCTGGCGCCCAATA 1080
QY 1081 CGAAAACCGCTCTCCCGCGGTTGGCGGATTCATTATGAGCTGGGACGACGAGGTTT 1140
DB 1081 CGAAAACCGCTCTCCCGCGGTTGGCGGATTCATTATGAGCTGGGACGACGAGGTTT 1140
QY 1141 CCGACTTGAAGAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCAATTAG 1200
DB 1141 CCGACTTGAAGAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCAATTAG 1200
QY 1201 GCACAAATTCATGTTTGCAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260
DB 1201 GCACAAATTCATGTTTGCAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGGCGAGCCATCGGAAGCTGTGGTATGGCTGTGCAAGTCGTAATAATCACTGCATAATTGC 1320
DB 1261 TCAGGCGAGCCATCGGAAGCTGTGGTATGGCTGTGCAAGTCGTAATAATCACTGCATAATTGC 1320

Qy 1321 TGTGGCTCAAGCGCGACCTCCGTTCTGGATTAATGTTTTTTGGCGCGACATCATACGGTT 1380
Db 1321 TGTGGCTCAAGCGCGACCTCCGTTCTGGATTAATGTTTTTTGGCGCGACATCATACGGTT 1380
Qy 1381 CTGGCAAAATATCTTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAAAATATCTTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
Qy 1441 ATTGTAGCGGATAAACAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCAGGA 1500
Db 1441 ATTGTAGCGGATAAACAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTGTTTTCAGGA 1500
Qy 1501 GCATTTACCAACAGGACCATAGATTATGAACCTGAAGAGTAAACTGGTAATCTGG 1560
Db 1501 GCATTTACCAACAGGACCATAGATTATGAACCTGAAGAGTAAACTGGTAATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATAACGGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAACTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGATAAACTGGAAGAGAAATTCACAGGTT 1680
Qy 1681 CGGCAACTGGCGATGGCCCTGACATTATCTTGGGCAACAGCCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGGCGATGGCCCTGACATTATCTTGGGCAACAGCCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCTGTTCGCTGAAATCACCCGGAACAGCGTTCCAGGACAGCTGAT 1800
Db 1741 GCTCAATCTGGCCTGTTCGCTGAAATCACCCGGAACAGCGTTCCAGGACAGCTGAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTTACCAACGGCAGCTGATGCTTTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACCAACGGCAGCTGATGCTTTTACCCGATCGCTGT 1860
Qy 1861 GAAGGTTATCGTGATTTATAAAGATCTGCTGCCGACACCCGCCAAACCTGGGAA 1920
Db 1861 GAAGGTTATCGTGATTTATAAAGATCTGCTGCCGACACCCGCCAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATTAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTCGCGCTGATGCTGCTGACGGGGTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTCGCGCTGATGCTGCTGACGGGGTTATCGGTTCAAG 2040
Qy 2041 TATGAAAACGGCAAGTACGACATTAAGACAGTGGCGTGGATACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACAGTGGCGTGGATACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGGTTGACCTGATTAAGAAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAAGAAACACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCTTAAATAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTAAATAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
Qy 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTAATGTTGTAACGCTGCTGCCGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTAATGTTGTAACGCTGCTGCCGACCTTC 2280
Qy 2281 AAGGTCACACCTCAAAACCGTTGCTGGCGTGGAGGAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGTCACACCTCAAAACCGTTGCTGGCGTGGAGGAGGTATTAACGCCGCCAGT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACATGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGACATGATGAAGTCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCCGCTGGGTGCCGTAGCGCTGAAGTCTTTACGAGGAAG 2460

Db 2401 GAAGCGGTTAATAAGACAAACCCGCTGGGTGCCGTAGCGCTGAAGTCTTTACGAGGAAG 2460
Qy 2461 TTGCGCAAGAGATCCACGATTTTGGCGCCACCATGGAAGAAACGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGCGCAAGAGATCCACGATTTTGGCGCCACCATGGAAGAAACGCCAGAAAGGTGAAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCGCTTTCTGGTATCGGTCGCTGCTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCGCTTTCTGGTATCGGTCGCTGCTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGCTGTCGATGAAGCCCTGAAAGACGCGCAGACTTAATTCGAGCTCG 2640
Qy 2641 AACCAACCAACAATAACAATAACAACAACCTCGGGATCGAGGGAAGGATTCAGAATTC 2700
Db 2641 AACCAACCAACAATAACAATAACAACAACCTCGGGATCGAGGGAAGGATTCAGAATTC 2700
RESULT 10
ADO23603
ID ADO23603 standard; DNA; 7352 BP.
XX ADO23603;
XX 01-JUL-2004 (first entry)
XX DNA encoding MBP-ToxoP30del14C(52-294aa) fusion protein.
XX P30 antigen; Toxo30del13C; Toxo30del12C; ToxoP30 MIX1;
KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
KW Toxoplasma gondii; ds; gene.
XX Toxoplasma gondii.
OS Synthetic.
XX US2004067239-A1.
XX 08-APR-2004.
PF 02-OCT-2002; 2002US-00263153.
XX 02-OCT-2002; 2002US-00263153.
PR (MAIN/) MAINE G T.
PA (PATE/) PATEL C B.
PA (GINS/) GINSBURG S R.
PA (BLIE/) BLIESE T R.
XX Maine GT, Patel CB, Gineburg SR, Bliese TR;
XX WPI; 2004-304563/28.
P-PSDB; ADO23604.
XX Novel purified polypeptide having sequence identity to amino acid sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C, Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG antibodies to Toxoplasma gondii.
XX Example 2; Fig 19; 114pp; English.
XX The invention relates to a purified P30 antigen (I) chosen from 3 fully defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6 amino acids added to the C-terminus of the amino acid sequence of Toxo30del13C P30 antigen sequence, or comprising an amino acid sequence in which at least one of the five C-terminal cysteine amino acids of the amino acid sequence of Toxo30del13C P30 antigen sequence is substituted with alanine, or comprising the amino acid sequence chosen from MBP-Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1 fusion proteins. (I) is useful for detecting the presence of IgM antibodies to Toxoplasma gondii in a test sample, which involves contacting the test sample suspected of containing the IgM antibodies

CC with a composition comprising (I) and detecting the presence of (I)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
XX Toxop30 fusion protein of the invention.

SQ Sequence 7352 BP; 1896 A; 1852 C; 1943 G; 1661 T; 0 U; 0 Other;

Query Match	81.5%;	Score 2690.4;	DB 12;	Length 7352;
Best Local Similarity	99.8%;	Pred. No. 0;		
Matches 2694;	Conservative	0;	Mismatches 6;	Indels 0; Gaps 0;

QY	1	CGGACCAATCGAATGGTGCAGAAACCTTTGCGGATATGGCATGATGAGCCCGGAGAGA	60
DB	1	CGGACCAATCGAATGGTGCAGAAACCTTTGCGGATATGGCATGATGAGCCCGGAGAGA	60
QY	61	GTCAATTCAGGGTGGTGAATGTGAACCAAGTAAAGTATACGATGTCGACAGTATGCGG	120
DB	61	GTCAATTCAGGGTGGTGAATGTGAACCAAGTAAAGTATACGATGTCGACAGTATGCGG	120
QY	121	GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACCAAGCGGATGATGATGTCGAAAA	180
DB	121	GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACCAAGCGGATGATGATGTCGAAAA	180
QY	181	CGCGGAAAGTGAAGCGGCGATGCGGAGCTGAATTAATCCCAACCGGTGGAC	240
DB	181	CGCGGAAAGTGAAGCGGCGATGCGGAGCTGAATTAATCCCAACCGGTGGAC	240
QY	241	AACAACTGGCGGCAAAACAGTCTGTGCTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGC	300
DB	241	AACAACTGGCGGCAAAACAGTCTGTGCTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGC	300
QY	301	ACGGCCGTGCGAATTTGTCGGGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360
DB	301	ACGGCCGTGCGAATTTGTCGGGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG	360
QY	361	TGGTGGTCTCGATGTAAGCGGCGATGCGGAGCTGTAATTAATCGGCGGATCAATCGG	420
DB	361	TGGTGGTCTCGATGTAAGCGGCGATGCGGAGCTGTAATTAATCGGCGGATCAATCGG	420
QY	421	TTCTCGGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACGAGATGCCA	480
DB	421	TTCTCGGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGGATGACGAGATGCCA	480
QY	481	TTGCTGTGGAAGCTGCGCTGCACTAATGTTCCGGGCTTAATTTCTTGATGTTCTGACCA	540
DB	481	TTGCTGTGGAAGCTGCGCTGCACTAATGTTCCGGGCTTAATTTCTTGATGTTCTGACCA	540
QY	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGACTGGGCTGGAGCATC	600
DB	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGACTGGGCTGGAGCATC	600
QY	601	TGGTCGATTTGGGTACAGCAATTCGGTGTAGCGGGCCCATTAAGTTCTGTCGCG	660
DB	601	TGGTCGATTTGGGTACAGCAATTCGGTGTAGCGGGCCCATTAAGTTCTGTCGCG	660
QY	661	CGGCTCTGCGTCTGGCTGGGATTAATCTCACTCGCAATCAATTCAGCCGATAG	720
DB	661	CGGCTCTGCGTCTGGCTGGGATTAATCTCACTCGCAATCAATTCAGCCGATAG	720
QY	721	CGGAAACGGGAGGCGATGAGTGCATGTCGGTGTTCGCAACCATGCAATGCTGA	780
DB	721	CGGAAACGGGAGGCGATGAGTGCATGTCGGTGTTCGCAACCATGCAATGCTGA	780
QY	781	ATGAGGCGATGTTCCCACTCGGATGTCGGTGTTCGCAACCATGCAATGCTGA	840
DB	781	ATGAGGCGATGTTCCCACTCGGATGTCGGTGTTCGCAACCATGCAATGCTGA	840
QY	841	TGGCGCGATTAACGAGTCCGGGCTGGGATGTCGGTGTTCGCAACCATGCAATGCTGA	900
DB	841	TGGCGCGATTAACGAGTCCGGGCTGGGATGTCGGTGTTCGCAACCATGCAATGCTGA	900
QY	901	ACGATACCGAAGACAGTCTATGTTATATCCCGCGTTTAAACCATCAACAGGATTTTC	960
DB	901	ACGATACCGAAGACAGTCTATGTTATATCCCGCGTTTAAACCATCAACAGGATTTTC	960

QY	961	GCCTGCTGGGCAAAACAGCGTGACCGCTTGTGCAACTCTCTCAGGCCAGCGGTGA	1020
DB	961	GCCTGCTGGGCAAAACAGCGTGACCGCTTGTGCAACTCTCTCAGGCCAGCGGTGA	1020
QY	1021	AGGCAATCAGCTGTCGCCGTCTCAGTGTGAAAGAAACACACCTTGGCGCCCAATA	1080
DB	1021	AGGCAATCAGCTGTCGCCGTCTCAGTGTGAAAGAAACACACCTTGGCGCCCAATA	1080
QY	1081	CGCAAAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGACAGTGGCAGCAGGTTT	1140
DB	1081	CGCAAAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGACAGTGGCAGCAGGTTT	1140
QY	1141	CCGACCTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTAGTTAGTCACTATTAG	1200
DB	1141	CCGACCTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTAGTTAGTCACTATTAG	1200
QY	1201	GCACAAATCTCATGTTTGCAGAGCTTATCATGCTGACGTCAGCGTGCACCAATGCTTCTGGCG	1260
DB	1201	GCACAAATCTCATGTTTGCAGAGCTTATCATGCTGACGTCAGCGTGCACCAATGCTTCTGGCG	1260
QY	1261	TCAGGACGCAATCGGAAGCTGTGTATGCTGTGAGTGTGCTGAGTGTGTAATCACTGCATAAATTCG	1320
DB	1261	TCAGGACGCAATCGGAAGCTGTGTATGCTGTGAGTGTGTAATCACTGCATAAATTCG	1320
QY	1321	TGTCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTTGGCGCGACATCAAAACGGTT	1380
DB	1321	TGTCGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTTGGCGCGACATCAAAACGGTT	1380
QY	1381	CTGCAAAATTTCTGAAATGAGCTGTGCAAAATTAATCATCGGCTCGTAAATGTTGGGA	1440
DB	1381	CTGCAAAATTTCTGAAATGAGCTGTGCAAAATTAATCATCGGCTCGTAAATGTTGGGA	1440
QY	1441	ATTGTGAGCGGATTAACAATTTCAACAGGAAACAGCGAGTCCGTTTGGTGTTCACGA	1500
DB	1441	ATTGTGAGCGGATTAACAATTTCAACAGGAAACAGCGAGTCCGTTTGGTGTTCACGA	1500
QY	1501	GCATTTCAACCAACAGGACCATAGATTTGAAATCTGAAAGCTGAAAGTAACTGGTAACTCG	1560
DB	1501	GCATTTCAACCAACAGGACCATAGATTTGAAATCTGAAAGCTGAAAGTAACTGGTAACTCG	1560
QY	1561	ATTAAACGGGATTAAGGCTATTAACAGGCTATTAACAGGCTATTAACAGGCTATTAACAGG	1620
DB	1561	ATTAAACGGGATTAAGGCTATTAACAGGCTATTAACAGGCTATTAACAGGCTATTAACAGG	1620
QY	1621	ACCGGAATTAAGTCAACGCTTGAAGTATCGGATTAACGAGGAAATTCGCAAGGTT	1680
DB	1621	ACCGGAATTAAGTCAACGCTTGAAGTATCGGATTAACGAGGAAATTCGCAAGGTT	1680
QY	1681	CGGCAACTGGCGATGCGCTGACATTTCTCTGGGCAACGACCGCTTGGTGGCTAC	1740
DB	1681	CGGCAACTGGCGATGCGCTGACATTTCTCTGGGCAACGACCGCTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCAGGACAAAGCTGTAT	1800
DB	1741	GCTCAATCTGGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCAGGACAAAGCTGTAT	1800
QY	1801	CGGTTTACCTGGATGCGCTGACATTAACAGGCAAGCTGATTCGCTTACCGGATCGCTGTT	1860
DB	1801	CGGTTTACCTGGATGCGCTGACATTAACAGGCAAGCTGATTCGCTTACCGGATCGCTGTT	1860
QY	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCAAACCCGCAAAACCTCGGAA	1920
DB	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCAAACCCGCAAAACCTCGGAA	1920
QY	1921	GAGATCCCGCGCTGGATTAAGAACTGAAAGGAAAGGTAAGAGCGCTGATGTTTCAAC	1980
DB	1921	GAGATCCCGCGCTGGATTAAGAACTGAAAGGAAAGGTAAGAGCGCTGATGTTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTTCACCTGGCGCTGATGTTGCTGACGCGGTTATCGGTTCAAG	2040
DB	1981	CTGCAAGAACCGTACTTTCACCTGGCGCTGATGTTGCTGACGCGGTTATCGGTTCAAG	2040

Db 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCACTGGCGCGTGGAGCATC 600
Qy 601 TGGTCGCATTGGGTCCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Db 601 TGGTCGCATTGGGTCCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGTCTCGG 660
Qy 661 CGCGTCTCGCTCTGGCTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db 661 CGCGTCTCGCTCTGGCTGGCTGGCAATAAATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Qy 721 CGGAACGGGAAGGACATGGAGTGCATATCTCGGTTTTCACAAACCAATGCAATTCGTA 780
Db 721 CGGAACGGGAAGGACATGGAGTGCATATCTCGGTTTTCACAAACCAATGCAATTCGTA 780
Qy 781 ATGAGGGCATCTGTTCCCACTCGCATGCTGTGTCACAGATCAGATGCGCTGGCGGCAA 840
Db 781 ATGAGGGCATCTGTTCCCACTCGCATGCTGTGTCACAGATCAGATGCGCTGGCGGCAA 840
Qy 841 TGCAGGCAATTCACGAGTCCGGGCTGGCGTGTGGTGGGATATCTCGTGTAGTGGGATACG 900
Db 841 TGCAGGCAATTCACGAGTCCGGGCTGGCGTGTGGTGGGATATCTCGTGTAGTGGGATACG 900
Qy 901 ACAGATCCGGAAGACGTCTATTTATATCCCGCGGTTTAAACACATCAAAACAGGATTTTC 960
Db 901 ACAGATCCGGAAGACGTCTATTTATATCCCGCGGTTTAAACACATCAAAACAGGATTTTC 960
Qy 961 GCCTGTCTGGGGCAAAACAGCGTGCACCGCTTGTCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGTCTGGGGCAAAACAGCGTGCACCGCTTGTCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Qy 1021 AGGCAATCAGCTGTGTCGCTCTCACTGTGTGAAAGAAAACCAACCTCGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTGTCGCTCTCACTGTGTGAAAGAAAACCAACCTCGCGCCCAATA 1080
Qy 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTAATGTAGTTAGCTCAGCTCATTATAG 1140
Db 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGGATTAATGTAGTTAGCTCAGCTCATTATAG 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGAAACGCAATTAATGTAGTTAGCTCAGCTCATTATAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGAAACGCAATTAATGTAGTTAGCTCAGCTCATTATAG 1200
Qy 1201 GCACAAATCTCATGTTTGACAGCTTATCATGACTGCAAGCGTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATGACTGCAAGCGTGACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGAGCCATCGGAAGCTGTGTATGGCTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGTATGGCTGTGCAAGTGTGCAAGTGTGCAAGTGTGCAAG 1320
Qy 1321 TGTGCTCAAGGGCACTCCCGTTCTGGATAATGTTTTTTCGCGGACATCATAAACGGTT 1380
Db 1321 TGTGCTCAAGGGCACTCCCGTTCTGGATAATGTTTTTTCGCGGACATCATAAACGGTT 1380
Qy 1381 CTGGCAATATCTGAATGAGCTGTGTGACAAATTAATCATCGGCTCGTATTAATGTGTGGA 1440
Db 1381 CTGGCAATATCTGAATGAGCTGTGTGACAAATTAATCATCGGCTCGTATTAATGTGTGGA 1440
Qy 1441 ATTGTAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTATAGGTGTTTTTACGA 1500
Db 1441 ATTGTAGCGGATAACAAATTTTCAACAGGAAACAGCCAGTCCGTTTATAGGTGTTTTTACGA 1500
Qy 1501 GCACCTTCAACAAAGGACATGATATTAAGAACTGAAAGGTAACCTGGTAAATCTCG 1560
Db 1501 GCACCTTCAACAAAGGACATGATATTAAGAACTGAAAGGTAACCTGGTAAATCTCG 1560
Qy 1561 ATTAAGCGGATTAAGGCTATACGCTCTCGTGAAGTCCGTAAGAAATTCGGAAGAT 1620
Db 1561 ATTAAGCGGATTAAGGCTATACGCTCTCGTGAAGTCCGTAAGAAATTCGGAAGAT 1620
Qy 1621 ACCGGAATTAAGTTCACCGTTGAGCATCCGGATAAATCGGAAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTTCACCGTTGAGCATCCGGATAAATCGGAAGAAATTCACAGGTT 1680

Qy 1681 GCGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACAGCCCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCAACAGCCCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTGGTGGAAATCAACCCCGACAAAGGTTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGTGGAAATCAACCCCGACAAAGGTTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCCGTACGTTTAAACAGGCAAGCTGTATTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCCGTACGTTTAAACAGGCAAGCTGTATTTACCCGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTCGCAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTCGCAACCCGCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTTTCACTGGCGCTGATTTGCTGCTGACCGGGGTTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTTCACTGGCGCTGATTTGCTGCTGACCGGGGTTTATGCGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTTAAAGAGCTGGCGTGGATAAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTTAAAGAGCTGGCGTGGATAAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGTTGACCTGATTTAAAAACAAACACATGAATGCAGACACCGGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTTAAAAACAAACACATGAATGCAGACACCGGATTAC 2160
Qy 2161 TCCATCGCAGAACCTGCTTTTAAAGGCGAAACAGCGATGACCATTAACGCGCGGTGG 2220
Db 2161 TCCATCGCAGAACCTGCTTTTAAAGGCGAAACAGCGATGACCATTAACGCGCGGTGG 2220
Qy 2221 GCATGTGCAACATCCAGACACAGCAAGTGAATTTATGGTGTAAACGCTTCTGCGGACCTTC 2280
Db 2221 GCATGTGCAACATCCAGACACAGCAAGTGAATTTATGGTGTAAACGCTTCTGCGGACCTTC 2280
Qy 2281 AAGGGTCAACCATCCAAACCGTTTGGCTGCTGAGCGCGAGGTATTAAACGCGCGCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTTGGCTGCTGAGCGCGAGGTATTAAACGCGCGCAGT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
Qy 2401 GAAGCGGTTTAAATAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTTAAATAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGGGAAAGATCCACGTAATTCGCCCAACATGGAACAAACCGCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGGAAAGATCCACGTAATTCGCCCAACATGGAACAAACCGCCAGAAAGGTGAAATCATG 2520
Qy 2521 CCGAACATCCCGAGATGTCGCTTCTGGTATGCCGTGCGTACTCGGTTGATCAAGCC 2580
Db 2521 CCGAACATCCCGAGATGTCGCTTCTGGTATGCCGTGCGTACTCGGTTGATCAAGCC 2580
Qy 2581 GCCAGCGCTCGTCAGACTGTCGATGAAGCCCTGAAAGAGCGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGCTCGTCAGACTGTCGATGAAGCCCTGAAAGAGCGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACACAAACAACTAACTAACTAACTAACTAACTAACTAACTAACTAACTAACTAACTAACT 2700
Db 2641 AACACAAACAACTAACTAACTAACTAACTAACTAACTAACTAACTAACTAACTAACT 2700

XX AC ADO23639;
 XX DT 01-JUL-2004 (first entry)
 XX DE DNA encoding MBP-ToxoP30MIX1 fusion protein.
 XX KW P30 antigen; Toxo30del13C; Toxo30del2C; ToxoP30 MIX1;
 KW MBP-Toxo30del13C(52-300aa); MBP-Toxo30del14C(52-294aa); MBP-Toxo30MIX1;
 KW Toxoplasma gondii; ds; gene.
 XX OS Toxoplasma gondii.
 OS Synthetic.
 XX PN US2004067239-A1.
 XX PD 08-APR-2004.
 XX PF 02-OCT-2002; 2002US-00263153.
 XX PR 02-OCT-2002; 2002US-00263153.
 XX PA (MAIN/) MAINE G T.
 PA (PATE/) PATEL C B.
 PA (GINS/) GINSBURG S R.
 PA (BLIE/) BLIESE T R.
 XX FI Maine GT, Patel CB, Ginsburg SR, Bliese TR;
 XX WPI; 2004-304563/28.
 DR P-P8DB; ADO23640.
 XX PT Novel purified polypeptide having sequence identity to amino acid
 PT sequence of P30 antigen of Toxoplasma gondii e.g. Toxo30del13C,
 PT Toxo30del12C, ToxoP30 MIX1, useful for detecting presence of IgM or IgG
 PT antibodies to Toxoplasma gondii.
 XX Example 5; Fig 31; 114pp; English.
 XX The invention relates to a purified P30 antigen (I) chosen from 3 fully
 CC defined Toxo30del13C, Toxo30del12C and ToxoP30 MIX1 sequences, having 1-6
 CC amino acids added to the C-terminus of the amino acid sequence of
 CC Toxo30del12C P30 antigen sequence, or comprising an amino acid sequence in
 CC which at least one of the five C-terminal cysteine amino acids of the
 CC amino acid sequence of Toxo30del13C P30 antigen sequence is substituted
 CC with alanine, or comprising the amino acid sequence chosen from MBP-
 CC Toxo30del13C(52-300aa), MBP-Toxo30del14C(52-294aa) and MBP-Toxo30MIX1
 CC fusion proteins. (I) is useful for detecting the presence of IgM
 CC antibodies to Toxoplasma gondii in a test sample, which involves
 CC contacting the test sample suspected of containing the IgM antibodies
 CC with a composition comprising (I) and detecting the presence of (I)/IgM
 CC antibody complexes. The present sequence represents DNA encoding a MBP-
 CC ToxoP30 fusion protein of the invention.
 XX SQ Sequence 7370 BP; 1900 A; 1856 C; 1950 G; 1664 T; 0 U; 0 Other;
 Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
 DB 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
 QY 61 GTCATTCAGGTGGTGAATGTGAACAGTAACGTTATACGATCTCCAGAGTATGCG 120
 DB 61 GTCATTCAGGTGGTGAATGTGAACAGTAACGTTATACGATCTCCAGAGTATGCG 120
 QY 121 GTGTCTCTTATCAGACCGGTTTCCCGGTGGTGAACAGGCCAGCCACGTTTCTGCGAAAA 180
 DB 121 GTGTCTCTTATCAGACCGGTTTCCCGGTGGTGAACAGGCCAGCCACGTTTCTGCGAAAA 180
 QY 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC 240

DB 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGGAGCTGAATTACATTTCCCAACCGCGTGGCAC 240
 QY 241 AACAACTGGCGGCAAAACAGTCTGTTGCTGATTTGGCGTTGGCCACCTCCAGTCTGGGCCCTGC 300
 DB 241 AACAACTGGCGGCAAAACAGTCTGTTGCTGATTTGGCGTTGGCCACCTCCAGTCTGGGCCCTGC 300
 QY 301 ACGCGCGTCCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTCGGTGGTCCAGCG 360
 DB 301 ACGCGCGTCCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTCGGTGGTCCAGCG 360
 QY 361 TGGTGGTGTGATGATAGAAAGAGCGCGCTGGAAGCTGTAAAGCGCGGTGCAAAATC 420
 DB 361 TGGTGGTGTGATGATAGAAAGAGCGCGCTGGAAGCTGTAAAGCGCGGTGCAAAATC 420
 QY 421 TTCTGCGCAACGCGTCACTGAGTGGGCTGATCATTAACATATCCGCTGATGACCAAGTATGCA 480
 DB 421 TTCTGCGCAACGCGTCACTGAGTGGGCTGATCATTAACATATCCGCTGATGACCAAGTATGCA 480
 QY 481 TTGCTGTGGAAGTGCCTGCACTAATGTTCCGGCGTTAATTTCTTGATGTCTGACACAGA 540
 DB 481 TTGCTGTGGAAGTGCCTGCACTAATGTTCCGGCGTTAATTTCTTGATGTCTGACACAGA 540
 QY 541 CACCCATCAACAGTATTAATTTTCTCCCATGAAGACGCTACGCACTGGCGGTGGAGCATC 600
 DB 541 CACCCATCAACAGTATTAATTTTCTCCCATGAAGACGCTACGCACTGGCGGTGGAGCATC 600
 QY 601 TGGTGCATTTGGGTCAACAGCAAAATTCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCTCGG 660
 DB 601 TGGTGCATTTGGGTCAACAGCAAAATTCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCTCGG 660
 QY 661 CGCGTCTGCGTCTGGCTGGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
 DB 661 CGCGTCTGCGTCTGGCTGGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
 QY 721 CGGAACGGGAAGGACGTGGAGTGCATGTCGGGTTTCAACAAACCATGCAAAATGCTGA 780
 DB 721 CGGAACGGGAAGGACGTGGAGTGCATGTCGGGTTTCAACAAACCATGCAAAATGCTGA 780
 QY 781 ATGAGGGCATGTTCCCACTCGCATGCTGGTTGTCACACGATCAGATGGCGCTGGCGGCAAA 840
 DB 781 ATGAGGGCATGTTCCCACTCGCATGCTGGTTGTCACACGATCAGATGGCGCTGGCGGCAAA 840
 QY 841 TGGCGGCCATTAACGAGTCCGGGCTGGCGGTGGTGGTGGGATATCTCGGTAGTGGGATAG 900
 DB 841 TGGCGGCCATTAACGAGTCCGGGCTGGCGGTGGTGGTGGGATATCTCGGTAGTGGGATAG 900
 QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACACCATCAACAGGATTTTC 960
 DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGGTTAACACCATCAACAGGATTTTC 960
 QY 961 GCCTGCTGGGCAAAACACGAGCGTGGACCGCTTGTGTCGAACCTCTCAGGGCCAGCGGCTGA 1020
 DB 961 GCCTGCTGGGCAAAACACGAGCGTGGACCGCTTGTGTCGAACCTCTCAGGGCCAGCGGCTGA 1020
 QY 1021 AGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAAAAACCCACCTGGCGCCCAATA 1080
 DB 1021 AGGCAATCAGCTGTTGCCCGCTCTCACTGGTGAAGAAAAAACCCACCTGGCGCCCAATA 1080
 QY 1081 CGAAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAATGCAAGCTGGCAACAGAGTTT 1140
 DB 1081 CGAAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAATGCAAGCTGGCAACAGAGTTT 1140
 QY 1141 CCGAGTCTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTACTCATTTAG 1200
 DB 1141 CCGAGTCTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCTACTCATTTAG 1200
 QY 1201 GCACAATTTCTCATGTTTTCACAGCTTTATCATGCACTGCAAGTGGTGCACCAATGTTCTGGCG 1260
 DB 1201 GCACAATTTCTCATGTTTTCACAGCTTTATCATGCACTGCAAGTGGTGCACCAATGTTCTGGCG 1260
 QY 1261 TCAGGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGTAATATCATGATTAATTCG 1320

CC antibodies to *Toxoplasma gondii* in a test sample, which involves
CC contacting the test sample suspected of containing the IgM antibodies
CC with a composition comprising (i) and detecting the presence of (i)/IgM
CC antibody complexes. The present sequence represents DNA encoding a MBP-
CC Toxop30 fusion protein of the invention.
XX
SQ

Sequence 7370 BP; 1900 A; 1857 C; 1950 G; 1663 T; 0 U; 0 Other;
Query Match 81.5%; Score 2690.4; DB 12; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY	1	CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATCGCATATGAGCCCGGAGAGA	60
Db	1	CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGCGATAGCGCCCGGAGAGA	60
QY	61	GTCAATTCAGGGTGTGAATGTGAACACAGTAACAGTTATACGATGTCGACAGTATGCCG	120
Db	61	GTCAATTCAGGGTGTGAATGTGAACACAGTAACAGTTATACGATGTCGACAGTATGCCG	120
QY	121	GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACACAGGCAGCCACGTTTCTGCGAAAA	180
Db	121	GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACACAGGCAGCCACGTTTCTGCGAAAA	180
QY	181	CGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTAACATCCCAACCGCGTGGCAC	240
Db	181	CGCGGAAAAAGTGAAGCGCGATGGCGAGCTGAATTAACATCCCAACCGCGTGGCAC	240
QY	241	AACAACTCGCGGGGAAAAAGTGGTGTGATGGGGTTGCCACCTTCAGTCTGGGCCCTGC	300
Db	241	AACAACTCGCGGGGAAAAAGTGGTGTGATGGGGTTGCCACCTTCAGTCTGGGCCCTGC	300
QY	301	ACGGCCGTTCGCAATGTCCGGCGGATTAATCTTCGCGCGATCAATGGGTGCCAGCG	360
Db	301	ACGGCCGTTCGCAATGTCCGGCGGATTAATCTTCGCGCGATCAATGGGTGCCAGCG	360
QY	361	TGGTGTTCGATGCTAGACGAGCGGTGGAAGCTGTGAAGCGGGGTGCAACATC	420
Db	361	TGGTGTTCGATGCTAGACGAGCGGTGGAAGCTGTGAAGCGGGGTGCAACATC	420
QY	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTTCGCTGATGATCAACAGGATGCCA	480
Db	421	TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTTCGCTGATGATCAACAGGATGCCA	480
QY	481	TGCTGTGGAAGCTCCCTGCAATAATGTTCGGCGTATTTCTTGATGTCTCTGACCAGA	540
Db	481	TGCTGTGGAAGCTCCCTGCAATAATGTTCGGCGTATTTCTTGATGTCTCTGACCAGA	540
QY	541	CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGGACTGGCGGTGGAGCATC	600
Db	541	CACCCATCAACAGTATTATTTCTCCCATGAAGACGGTACCGGACTGGCGGTGGAGCATC	600
QY	601	TGGTGGCATTTGGTCCACGAAATTCGCGCTGTAGCGGGCCATTAAGTTCTGTCTCGG	660
Db	601	TGGTGGCATTTGGTCCACGAAATTCGCGCTGTAGCGGGCCATTAAGTTCTGTCTCGG	660
QY	661	CAGCTCTCGTCTGGCTGGCTGGAATAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Db	661	CAGCTCTCGTCTGGCTGGCTGGAATAATAATCTCACTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGAAACGGGAAGCGGACATGGAGTGCATGTCCGGTTTTCACAAACCATGCAATGCTGA	780
Db	721	CGAAACGGGAAGCGGACATGGAGTGCATGTCCGGTTTTCACAAACCATGCAATGCTGA	780
QY	781	ATGAGGGGATCGTCCCACTCGGATGCTGGTTGGCCACAGATCGGCTGGCGGCA	840
Db	781	ATGAGGGGATCGTCCCACTCGGATGCTGGTTGGCCACAGATCGGCTGGCGGCA	840
QY	841	TGCGCGCATTAACCGAGTCCGGGCTGCGGTGGTGGCGGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGCGCATTAACCGAGTCCGGGCTGCGGTGGTGGCGGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACCAACATCAACAGGATTTTC	960

Db	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACATCAACAGGATTTTC	960
QY	961	GCCTGCTGGGGCAAAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGCGGTGA	1020
Db	961	GCCTGCTGGGGCAAAACACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGCGGTGA	1020
QY	1021	AGGGCAATCAGCTGTGGCCCGTCTCATCTGGTGAAGAAAAACCAACCTTGGGGCCCAATA	1080
Db	1021	AGGGCAATCAGCTGTGGCCCGTCTCATCTGGTGAAGAAAAACCAACCTTGGGGCCCAATA	1080
QY	1081	CGAAAAACCGCTCTCCCGCGCTGGCGGATTCATTAATGAGCTGCGACGACAGGTTT	1140
Db	1081	CGAAAAACCGCTCTCCCGCGCTGGCGGATTCATTAATGAGCTGCGACGACAGGTTT	1140
QY	1141	CCGCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG	1200
Db	1141	CCGCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG	1200
QY	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGATGCGAGCTGCACCAATGCTTCTGGCG	1260
Db	1201	GCACAAATCTCATGTTTGACAGCTTATCATCGATGCGAGCTGCACCAATGCTTCTGGCG	1260
QY	1261	TCAGGACGCCATCGAAGCTGTGGTATGCGTGTGAGCTCGTAATCACTGCGATAATTCG	1320
Db	1261	TCAGGACGCCATCGAAGCTGTGGTATGCGTGTGAGCTCGTAATCACTGCGATAATTCG	1320
QY	1321	TGTCGCTCAAGCGCACCTCCCGTCTTGGAATAATGTTTTGCGCGGACATCAACCGTT	1380
Db	1321	TGTCGCTCAAGCGCACCTCCCGTCTTGGAATAATGTTTTGCGCGGACATCAACCGTT	1380
QY	1381	CTGGCAAAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTG	1440
Db	1381	CTGGCAAAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTG	1440
QY	1441	ATTGTGAGCGGATTAACAAATTTTACACAGGAAACAGCAGTCCGTTTAGGTGTTTTCACGA	1500
Db	1441	ATTGTGAGCGGATTAACAAATTTTACACAGGAAACAGCAGTCCGTTTAGGTGTTTTCACGA	1500
QY	1501	GCACCTTCAACCAACAGGACCATAGATTAAGAACTGAAGAGGTAATCTGTAATCTGG	1560
Db	1501	GCACCTTCAACCAACAGGACCATAGATTAAGAACTGAAGAGGTAATCTGTAATCTGG	1560
QY	1561	ATTAAACGGCGATAAAGGCTATAAAGCTCTCGCTGAGTTCGGTGAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAAACGGCGATAAAGGCTATAAAGCTCTCGCTGAGTTCGGTGAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAAGTACCGTTCAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT	1680
Db	1621	ACCGGAATTAAGTACCGTTCAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT	1680
QY	1681	CGGGCAACTGGCGATGGCCCTGACATTAATCTTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
Db	1681	CGGGCAACTGGCGATGGCCCTGACATTAATCTTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGGCGTGTGGCTGAAATACCCCGGCAAGAGCGTTCCAGACAAGCTGTAT	1800
Db	1741	GCTCAATCTGGCGTGTGGCTGAAATACCCCGGCAAGAGCGTTCCAGACAAGCTGTAT	1800
QY	1801	CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTGCTTACCGATCGCTGT	1860
Db	1801	CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTGCTTACCGATCGCTGT	1860
QY	1861	GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCGGAAACCCCGCAAAACCTTGGGAA	1920
Db	1861	GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCGGAAACCCCGCAAAACCTTGGGAA	1920
QY	1921	GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAGGTAAAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAGGTAAAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGGTTATCGGTTCAAG	2040

QY	541	CACCCATCAACAGTATTATTTCTCCCATGAGACGGTACGGACTGGCGCTGGAGCATC	600
Db	541	CACCCATCAACAGTATTATTTCTCCCATGAGACGGTACGGACTGGCGCTGGAGCATC	600
QY	601	TGGTCGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCGCG	660
Db	601	TGGTCGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCGCG	660
QY	661	CGCGTCTCGGTCTGGCTGGCTGGGATAAATATCTCATCTCGCAATCAAAATTCAGCCGATAG	720
Db	661	CGCGTCTCGGTCTGGCTGGGATAAATATCTCATCTCGCAATCAAAATTCAGCCGATAG	720
QY	721	CGGAACGGGAAGGGACCTGGAGTGCATGTCGGGTTTCAACAAACCATGCAAAATGCTGA	780
Db	721	CGGAACGGGAAGGGACCTGGAGTGCATGTCGGGTTTCAACAAACCATGCAAAATGCTGA	780
QY	781	ATGAGGGCATCGTTCCTCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGCGCGAA	840
Db	781	ATGAGGGCATCGTTCCTCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGCGCGAA	840
QY	841	TGCGGGCCATTACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGGGCCATTACGAGTCCGGGCTGCGGTTGGTGGCGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACCGGAAGACAGCTCATGTTATATCCCGCGTTAACCCACATCAACAGGATTTTC	960
Db	901	ACGATACCGGAAGACAGCTCATGTTATATCCCGCGTTAACCCACATCAACAGGATTTTC	960
QY	961	GCCTGCTGGGGCAACACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA	1020
Db	961	GCCTGCTGGGGCAACACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA	1020
QY	1021	AGGGCAATCAGTGTTGCCGCTCTCACTGGTGAAGAAAAACCCCTGGCGGCCAATA	1080
Db	1021	AGGGCAATCAGTGTTGCCGCTCTCACTGGTGAAGAAAAACCCCTGGCGGCCAATA	1080
QY	1081	CGCAAAACGCTCTCCCGCGGTTGGCGGATTCATTATGCGAGCTGGCAGCAGAGTTT	1140
Db	1081	CGCAAAACGCTCTCCCGCGGTTGGCGGATTCATTATGCGAGCTGGCAGCAGAGTTT	1140
QY	1141	CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG	1200
Db	1141	CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG	1200
QY	1201	GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGAGTGCACCAATGCTTCTGGCG	1260
Db	1201	GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGAGTGCACCAATGCTTCTGGCG	1260
QY	1261	TCAGGCAGCCATCGGAAGCTGTGGTATGCGTGTGAGGTTCGTAATCACTGCGATTAATTCG	1320
Db	1261	TCAGGCAGCCATCGGAAGCTGTGGTATGCGTGTGAGGTTCGTAATCACTGCGATTAATTCG	1320
QY	1321	TGTCGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGCGACATCAATACGGTT	1380
Db	1321	TGTCGCTCAAGCGCACTCCCGTCTCGGATATGTTTTTGGCGCGACATCAATACGGTT	1380
QY	1381	CTGGCAAAATTTCTGAAATAGAGCTGTGACATTAATCATCGGCTCGTATATGTTGGA	1440
Db	1381	CTGGCAAAATTTCTGAAATAGAGCTGTGACATTAATCATCGGCTCGTATATGTTGGA	1440
QY	1441	ATTGTGAGCGGATAACAAATTTCAACAGAAACAGCCAGTCCGTTTAGTGTTTTCAGGA	1500
Db	1441	ATTGTGAGCGGATAACAAATTTCAACAGAAACAGCCAGTCCGTTTAGTGTTTTCAGGA	1500
QY	1501	GCACCTTCCACCAAGGACCATAGATATGAAAACTGGAAGAGGTAACCTGATATCTGG	1560
Db	1501	GCACCTTCCACCAAGGACCATAGATATGAAAACTGGAAGAGGTAACCTGATATCTGG	1560
QY	1561	ATTAAACGGCGATAAAGGCTATAACCGTCTCGTGAAGTCGGTAAGAAAATTCAGAGAAAG	1620
Db	1561	ATTAAACGGCGATAAAGGCTATAACCGTCTCGTGAAGTCGGTAAGAAAATTCAGAGAAAG	1620
QY	1621	ACCGGAATTAAGTCACCGTTGAGCATCCGGATAAACTGGGAAGAGAAAATTTCCACAGGTT	1680

RESULT 15

Db	1621	ACCGGAATTAAGTCACCGTTGAGCATCCGGATTAACCTGGGAAGAGAAAATTTCCACAGGTT	1680
QY	1681	CGCGCAACTGCGCATGGCCCTTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
Db	1681	CGCGCAACTGCGCATGGCCCTTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC	1740
QY	1741	GCTCAATCTGCGCTGTGGCTGAAATCAACCCGGGCAAAAGGTTCCAGGACAAGCTGTAT	1800
Db	1741	GCTCAATCTGCGCTGTGGCTGAAATCAACCCGGGCAAAAGGTTCCAGGACAAGCTGTAT	1800
QY	1801	CGGTTTACCTGGGATGGCGTACGTTTACAAACGCAAGCTGATTGCTTACCCGATCGCTGT	1860
Db	1801	CGGTTTACCTGGGATGGCGTACGTTTACAAACGCAAGCTGATTGCTTACCCGATCGCTGT	1860
QY	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA	1920
Db	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAAAACCTGGGAA	1920
QY	1921	GAGATCCCGCGCTGGATTAAGAACTGAAGGAAAGTAAAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATTAAGAACTGAAGGAAAGTAAAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTCCACCTGGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTCCACCTGGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG	2040
QY	2041	TATGAAAAACGCAAGTACGACATTAAGAACGCTGGGCGTGGATAACGCTGGCGCGAAAGCG	2100
Db	2041	TATGAAAAACGCAAGTACGACATTAAGAACGCTGGGCGTGGATAACGCTGGCGCGAAAGCG	2100
QY	2101	GGTCTGACCTTCTCGTTTGACCTGATTAATAAACAACATGAATGCGAGACACCGATTAC	2160
Db	2101	GGTCTGACCTTCTCGTTTGACCTGATTAATAAACAACATGAATGCGAGACACCGATTAC	2160
QY	2161	TCCATCGCAGAGCTGCCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG	2220
Db	2161	TCCATCGCAGAGCTGCCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG	2220
QY	2221	GCATGGTCCAAACATCGACACACGCAAAAGTGAATTAATGGTGAACCGGTACTGCGGACCTTC	2280
Db	2221	GCATGGTCCAAACATCGACACACGCAAAAGTGAATTAATGGTGAACCGGTACTGCGGACCTTC	2280
QY	2281	AAGGTCAACCATCAAAACCGTTGCTGGCGTGTGAGCGAGGTATTAACCGCCCGCACT	2340
Db	2281	AAGGTCAACCATCAAAACCGTTGCTGGCGTGTGAGCGAGGTATTAACCGCCCGCACT	2340
QY	2341	CGGAACAAAGAGCTGGCAAAAGAGTTCTCGAANAACATATCTGCTGACTGATGAAGTCTG	2400
Db	2341	CGGAACAAAGAGCTGGCAAAAGAGTTCTCGAANAACATATCTGCTGACTGATGAAGTCTG	2400
QY	2401	GAAGCGGTTAATAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTTACGAGGAAGAG	2460
Db	2401	GAAGCGGTTAATAAGACAAACCGCTGGGTGCGTAGCGCTGAAGTCTTTACGAGGAAGAG	2460
QY	2461	TTGGCGAAAGATTCACGTTATTCGCGCCACCATGGAAGAAAACGCCAGAAAGGTGAATCATG	2520
Db	2461	TTGGCGAAAGATTCACGTTATTCGCGCCACCATGGAAGAAAACGCCAGAAAGGTGAATCATG	2520
QY	2521	CCGACATCCCGCAGATGTCGCTTCTTGATATGCGGTGCGTACTGCGGTGATCAACGCC	2580
Db	2521	CCGACATCCCGCAGATGTCGCTTCTTGATATGCGGTGCGTACTGCGGTGATCAACGCC	2580
QY	2581	GCCAGCGGTCTGACATGTCGATGAAGCCCTGGAAGACGCGCAGACTAAATTCGAGCTCG	2640
Db	2581	GCCAGCGGTCTGACATGTCGATGAAGCCCTGGAAGACGCGCAGACTAAATTCGAGCTCG	2640
QY	2641	AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGGAAGGATTTCAAGATTC	2700
Db	2641	AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGGAAGGATTTCAAGATTC	2700

QY 1261 TCAGGACGCATCGGAAGCTGGTATGGCTGTGCAGGTCGTAATCACTCATTAATTTCG 1320
Db 1261 TCAGGACGCATCGGAAGCTGGTATGGCTGTGCAGGTCGTAATCACTCATTAATTTCG 1320
QY 1321 TGTCCCTCAAGCGGCACATCCCGTCTCGGATATGTTTTTGGCCGACATCATACGGTT 1380
Db 1321 TGTCCCTCAAGCGGCACATCCCGTCTCGGATATGTTTTTGGCCGACATCATACGGTT 1380
QY 1381 CTGGCAATATTTCTGAATAGCTGTTGCAATTAATCATCGGCTCGTATATGTTGGA 1440
Db 1381 CTGGCAATATTTCTGAATAGCTGTTGCAATTAATCATCGGCTCGTATATGTTGGA 1440
QY 1441 ATTGTAGCGGATACAATTTTCAACAGCAACAGCCAGTCCGTTTAGTGTGTTTCACGA 1500
Db 1441 ATTGTAGCGGATACAATTTTCAACAGCAACAGCCAGTCCGTTTAGTGTGTTTCACGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTTATGAAAACTGAAGAGGTAATACTGGTAATCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTTATGAAAACTGAAGAGGTAATACTGGTAATCTGG 1560
QY 1561 ATTAAACGCGATAAAGGCTATTAACCGTCTCGTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAAACGCGATAAAGGCTATTAACCGTCTCGTGAAGTCGGTAAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATTAACCTGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTCACATTTATCTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTTATCTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTGGCTGAATTCACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTGGCTGAATTCACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTAGCTTACACGCAAGCTGATGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTAGCTTACACGCAAGCTGATGCTTACCCGATCGCTGTT 1860
QY 1861 GAAGGTTTATCGCTGATTTATAAAGAGATCTGCTGCCGACCCGCAAAACCTGGGAA 1920
Db 1861 GAAGGTTTATCGCTGATTTATAAAGAGATCTGCTGCCGACCCGCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAGCGAAAGGTAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGAGCGGGGTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGAGCGGGGTTATCGGTTCAAG 2040
QY 2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGCGTGGATAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGGCAAGTACGACATTAAGACGTGGCGTGGATAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTGTGACCTGATTAATAAACAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTGTGACCTGATTAATAAACAACACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAAAGCTGCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAAAGCTGCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCAACTCGACACCGCAAGTGAATTTAGTGTAAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAACTCGACACCGCAAGTGAATTTAGTGTAAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGGTCAACCATCCAAACCGTTGTTGGCTGCTGAGCGCAGGTATTAAACGCCGCCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTGTTGGCTGCTGAGCGCAGGTATTAAACGCCGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTTCCTCGAAAACTATCTGCTGATGAAGGTCTG 2400

Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTTCCTCGAAAACTATCTGCTGACTGATGAAGGTCTG 2400
QY 2401 GAAGCGGTTAATAAAGACAAAACCGCTGGGTGCGCTAGCGCTCAAGTCTTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAAACCGCTGGGTGCGCTAGCGCTCAAGTCTTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAAGATCCACGTTATTTGCCGCCACCAATGGAAAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAAGATCCACGTTATTTGCCGCCACCAATGGAAAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCGTGCCTGTAATCGCGTGCATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGCGTGCCTGTAATCGCGTGCATCAACGCC 2580
QY 2581 GCCAGCGTCTGTCAGACTGTGTCGATGAAGCCCTGAAAGAACGCCGACACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGTCAGACTGTGTCGATGAAGCCCTGAAAGAACGCCGACACTAAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAATTAACAATAACAACCTCGGGATCGAGGAGGATTTCAGAAATTC 2700
Db 2641 AACAAACAAACAATTAACAATAACAACCTCGGGATCGAGGAGGATTTCAGAAATTC 2700

Search completed: August 1, 2005, 04:28:52
Job time : 1102.6 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 352.4 Seconds
(without alignments)
15322.682 Million cell updates/sec

Title: US-09-765-555B-17
Perfect score: 3300
Sequence: 1 ccgacacacgaatgtgc.....acgacgttccggaactagct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818139359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2700	81.8	7475	2 US-08-971-036-1	Sequence 1, Appli
2	2700	81.8	7475	3 US-09-096-570-1	Sequence 1, Appli
3	2700	81.8	7475	3 US-09-265-617B-1	Sequence 1, Appli
c	4	1417.6	43.0	5926 3 US-09-027-169-3	Sequence 3, Appli
5	1362	41.3	5201	4 US-09-640-882-2	Sequence 2, Appli
6	1362	41.3	5201	4 US-09-640-882-3	Sequence 3, Appli
7	1204	36.5	4557	4 US-08-778-717-5	Sequence 5, Appli
8	1200.8	36.4	3832	1 US-08-148-675A-2	Sequence 2, Appli
c	9	1199.2	36.3	5248 3 US-08-487-283A-18	Sequence 18, Appli
c	10	1199.2	36.3	5248 5 PCT-US96-05611A-21	Sequence 21, Appli
c	11	1199.2	36.3	5312 4 US-10-263-103-35	Sequence 35, Appli
c	12	1199.2	36.3	5443 2 US-08-929-967-1	Sequence 1, Appli
c	13	1199.2	36.3	5502 4 US-09-702-705-785	Sequence 785, App
c	14	1199.2	36.3	5502 4 US-09-736-457-785	Sequence 785, App
c	15	1199.2	36.3	5502 4 US-09-614-124B-785	Sequence 785, App
c	16	1199.2	36.3	5502 4 US-09-671-325-785	Sequence 785, App
c	17	1199.2	36.3	5502 4 US-09-589-184-785	Sequence 785, App
c	18	1199.2	36.3	5502 4 US-09-658-824-785	Sequence 785, App
c	19	1199.2	36.3	5516 2 US-08-929-967-3	Sequence 3, Appli
c	20	1199.2	36.3	5873 4 US-09-695-437A-62	Sequence 62, Appli
c	21	1199.2	36.3	6353 4 US-09-702-705-784	Sequence 784, App
c	22	1199.2	36.3	6353 4 US-09-736-457-784	Sequence 784, App
c	23	1199.2	36.3	6353 4 US-09-614-124B-784	Sequence 784, App
c	24	1199.2	36.3	6353 4 US-09-671-325-784	Sequence 784, App
c	25	1199.2	36.3	6353 4 US-09-589-184-784	Sequence 784, App
c	26	1199.2	36.3	6353 4 US-09-658-824-784	Sequence 784, App
c	27	1199.2	36.3	6361 3 US-08-646-538-7	Sequence 7, Appli

c	28	1199.2	36.3	6361	3 US-09-503-222-7	Sequence 7, Appli
c	29	1199.2	36.3	6727	2 US-08-125-462-2	Sequence 2, Appli
c	30	1199.2	36.3	6727	2 US-08-891-848-2	Sequence 2, Appli
c	31	1199.2	36.3	6799	2 US-08-125-462-5	Sequence 5, Appli
c	32	1199.2	36.3	6799	2 US-08-891-848-5	Sequence 5, Appli
c	33	1199.2	36.3	7659	3 US-09-128-314-4	Sequence 4, Appli
c	34	1199.2	36.3	7659	3 US-09-056-556-213	Sequence 213, App
c	35	1199.2	36.3	7676	3 US-08-072-586-208	Sequence 208, App
c	36	1199.2	36.3	7676	4 US-09-072-967-213	Sequence 213, App
c	37	1199.2	36.3	7676	4 US-09-287-849-9	Sequence 9, Appli
c	38	1199.2	36.3	8031	3 US-09-643-597-254	Sequence 254, App
c	39	1199.2	36.3	8031	4 US-09-480-884A-254	Sequence 254, App
c	40	1199.2	36.3	8031	4 US-09-542-615A-254	Sequence 254, App
c	41	1199.2	36.3	8031	4 US-09-606-421B-254	Sequence 254, App
c	42	1199.2	36.3	8031	4 US-09-476-496A-254	Sequence 254, App
c	43	1199.2	36.3	8031	4 US-09-630-940B-254	Sequence 254, App
c	44	1199.2	36.3	8157	3 US-09-128-314-3	Sequence 3, Appli
c	45	1199.2	36.3	8501	3 US-08-793-900-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-08-971-036-1
; Sequence 1, Application US/08971036
; Patent No. 5866684
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R
; APPLICANT: Hurst, David N
; APPLICANT: Jones, Philip S
; APPLICANT: Kay, Paul B
; APPLICANT: Raynham, Tony M
; APPLICANT: Wilson, Francis X
; TITLE OF INVENTION: Amino Acid Derivatives
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingland Street
; CITY: Nutley
; STATE: N.J.
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/971.036
; FILING DATE: 14-NOV-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9623908.2
; FILING DATE: 18-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Kreisler, Lewis J
; REGISTRATION NUMBER: 38522
; REFERENCE/DOCKET NUMBER: RAN 4430/073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (973) 235-4387
; TELEFAX: (973) 235-2363
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7475 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
US-08-971-036-1

Query Match 81.8%; Score 2700; DB 2; Length 7475;

Best Local Similarity 100.0%; Pred. No. 0;			
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1	CGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGCGCATGATAGCCCGGAAGAGA	60
Db	1	CGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGCGCATGATAGCCCGGAAGAGA	60
QY	61	GTCAATTCAGGGTGGTGAATGTGAACACAGTAACCGTTATACGATGTCGCAGAGTATGCCG	120
Db	61	GTCAATTCAGGGTGGTGAATGTGAACACAGTAACCGTTATACGATGTCGCAGAGTATGCCG	120
QY	121	GTGTCCTTATCAGACCGTTTCCGGGTGGTGAACAGGCCAGCCAGCGTTTCTGCGAAAA	180
Db	121	GTGTCCTTATCAGACCGTTTCCGGGTGGTGAACAGGCCAGCCAGCGTTTCTGCGAAAA	180
QY	181	CCGCGGAAAAAGTGAAGCGGGGATGGCGAGCTGAATTAACATCCCAACCGGTGGCAC	240
Db	181	CCGCGGAAAAAGTGAAGCGGGGATGGCGAGCTGAATTAACATCCCAACCGGTGGCAC	240
QY	241	AACAATCGCGGCGAAAACAGTGGTGTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
Db	241	AACAATCGCGGCGAAAACAGTGGTGTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
QY	301	ACGCGCGTCCGCAATTCGCGGGGATTAATCTCGCGCGCATCAACTGGGTGCGACG	360
Db	301	ACGCGCGTCCGCAATTCGCGGGGATTAATCTCGCGCGCATCAACTGGGTGCGACG	360
QY	361	TGGTGGTTCGATGTTAGAACGAGCGGCTGAAAGCGCTGAAAGCGCGGTGCACAATC	420
Db	361	TGGTGGTTCGATGTTAGAACGAGCGGCTGAAAGCGCTGAAAGCGCGGTGCACAATC	420
QY	421	TTCTCGCGCAACGCGTCAAGTGGGTGATCAATTAATCTATCCGTCGATGACCAAGTCCCA	480
Db	421	TTCTCGCGCAACGCGTCAAGTGGGTGATCAATTAATCTATCCGTCGATGACCAAGTCCCA	480
QY	481	TGCTGTGCGAAGCTGCGTCAATGTTCCGGGTATTTCTTGATGTTCTCGACGAGA	540
Db	481	TGCTGTGCGAAGCTGCGTCAATGTTCCGGGTATTTCTTGATGTTCTCGACGAGA	540
QY	541	CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACCGACTGGCGGTGAGCATC	600
Db	541	CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACCGACTGGCGGTGAGCATC	600
QY	601	TGGTGCATTTGGTTCACAGCAATCGCGTGTAGCGGGCCCATTAAGTTCCTCTCGG	660
Db	601	TGGTGCATTTGGTTCACAGCAATCGCGTGTAGCGGGCCCATTAAGTTCCTCTCGG	660
QY	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCGGATAG	720
Db	661	CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATCAGCGGATAG	720
QY	721	CGGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTCAACAAACCATGCAATGCTGA	780
Db	721	CGGAACGGGAAGGCGACTGGAGTGCCATGTCCGGTTTCAACAAACCATGCAATGCTGA	780
QY	781	ATGAGGCGATGTTCCCACTCGCATGCTGGTTCGCAACGATCAGATGGCGTGGCGCAA	840
Db	781	ATGAGGCGATGTTCCCACTCGCATGCTGGTTCGCAACGATCAGATGGCGTGGCGCAA	840
QY	841	TGCGCGCATTTACCGAGTCCGGGTGGCGGTGGTGGCGATATCTCGGTAGTGGGATACG	900
Db	841	TGCGCGCATTTACCGAGTCCGGGTGGCGGTGGTGGCGATATCTCGGTAGTGGGATACG	900
QY	901	ACGATACGGAAGACAGTCTATTTATTCGCGCGTTTAAACACCATCAAAAGATTTTC	960
Db	901	ACGATACGGAAGACAGTCTATTTATTCGCGCGTTTAAACACCATCAAAAGATTTTC	960
QY	961	GCCTGTGGGGCAAAACAGGGTGGACCGTTGCTGCAACTCTCTCAGGCGACCGGTGA	1020
Db	961	GCCTGTGGGGCAAAACAGGGTGGACCGTTGCTGCAACTCTCTCAGGCGACCGGTGA	1020
QY	1021	AGGCAATCAGCTGTTCCCGCTCTCACTGGTGAAGAAAAACACCCCTGGCGCCCAATA	1080
Db	1021	AGGCAATCAGCTGTTCCCGCTCTCACTGGTGAAGAAAAACACCCCTGGCGCCCAATA	1080
QY	1081	CGAAAAACCGCTCTCCCGCGCGTGGCCGATTCATTAATGACAGTGGCAGCAGAGTTT	1140
Db	1081	CGAAAAACCGCTCTCCCGCGCGTGGCCGATTCATTAATGACAGTGGCAGCAGAGTTT	1140
QY	1141	CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCATTATAG	1200
Db	1141	CCCGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCATTATAG	1200
QY	1201	GCACAAATTCATGTTGACAGCTTATCATCGACTGACAGCTGACCAATGCTTCTCGCG	1260
Db	1201	GCACAAATTCATGTTGACAGCTTATCATCGACTGACAGCTGACCAATGCTTCTCGCG	1260
QY	1261	TCAGGAGCCATCGGAAGCTGTGTATGCTGTGACAGTCTGTAATCACTGCAATAATTCG	1320
Db	1261	TCAGGAGCCATCGGAAGCTGTGTATGCTGTGACAGTCTGTAATCACTGCAATAATTCG	1320
QY	1321	TGTCGCTCAAGGGCGACTCCCGTTCGGAATAATGTTTTTTCGCGCGACATCAACCGTT	1380
Db	1321	TGTCGCTCAAGGGCGACTCCCGTTCGGAATAATGTTTTTTCGCGCGACATCAACCGTT	1380
QY	1381	CTGGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCGCTCGTAAATGTTGGA	1440
Db	1381	CTGGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCGCTCGTAAATGTTGGA	1440
QY	1441	ATTGTGAGCGGATAACAAATTTCAACAGGAAACAGCCAGTCCGTTTAGTGTGTTTCA	1500
Db	1441	ATTGTGAGCGGATAACAAATTTCAACAGGAAACAGCCAGTCCGTTTAGTGTGTTTCA	1500
QY	1501	GCACCTTCAACCAACAGGACCATAGATTATGAAAGCTGAAAGGTAACCTGGTAATCTCG	1560
Db	1501	GCACCTTCAACCAACAGGACCATAGATTATGAAAGCTGAAAGGTAACCTGGTAATCTCG	1560
QY	1561	ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAAGTAAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAAACGGCGATAAAGGCTATAACGGTCTCGCTGAAAGTAAAGAAATTCGAGAAAGAT	1620
QY	1621	ACCGGAATTAAGCTACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT	1680
Db	1621	ACCGGAATTAAGCTACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT	1680
QY	1681	CGCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCAACAGCCGCTTGGTGCTAC	1740
Db	1681	CGCGCAACTGGCGATGGCCCTGACATTAATCTCTGGGCAACAGCCGCTTGGTGCTAC	1740
QY	1741	GCTCAATCTGGCTGTGGTGAATTAACCCCGGAACAAAGCGTTCCAGGACAGCTGTAT	1800
Db	1741	GCTCAATCTGGCTGTGGTGAATTAACCCCGGAACAAAGCGTTCCAGGACAGCTGTAT	1800
QY	1801	CCGTTTACCTGGGATGCGGTACGTTACAAACGGCAAGCTGATTCGTTACCGGATCGCTGT	1860
Db	1801	CCGTTTACCTGGGATGCGGTACGTTACAAACGGCAAGCTGATTCGTTACCGGATCGCTGT	1860
QY	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTCCGAAACCCGCAAAACCTCGGAA	1920
Db	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTCCGAAACCCGCAAAACCTCGGAA	1920
QY	1921	GAGATCCCGCGCTGATTAAGAACTGAAAGGTAAGAGCGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGCGCTGATTAAGAACTGAAAGGTAAGAGCGCGCTGATGTTCAAC	1980
QY	1981	CTGCAAGAACCGTACTTACCTGGCGCTGATTCGTTGACGGGGTTATGCGTTCAAG	2040
Db	1981	CTGCAAGAACCGTACTTACCTGGCGCTGATTCGTTGACGGGGTTATGCGTTCAAG	2040
QY	2041	TATGAAGAACCGCAAGTACGACATTAAGACGTTGGCGTGAATAACGCTGGCGGAAGCG	2100
Db	2041	TATGAAGAACCGCAAGTACGACATTAAGACGTTGGCGTGAATAACGCTGGCGGAAGCG	2100
QY	2101	GCTGTGACCTTCTGCTGACCTGATTAAGAAACCAACATGAATGCAGACACCGATTAC	2160
Db	2101	GCTGTGACCTTCTGCTGACCTGATTAAGAAACCAACATGAATGCAGACACCGATTAC	2160

2161	Qy	TCCATCGCAGAGCTGCTTTTAAATAAAGGGGAAACACGGATGACCATCAACCGGCCCTGTG	2222
2161	Db	TCCATCGCAGAAGCTGCTTTTAAATAAAGGGGAAACACGGATGACCATCAACCGGCCCTGTG	2220
2221	Qy	GCATGGTCCACATCGCACACGACCAAGTGAATTTATGGTGTAAAGGTACTGCCGACCTTC	2280
2221	Db	GCATGGTCCACATCGCACACGACCAAGTGAATTTATGGTGTAAAGGTACTGCCGACCTTC	2280
2281	Qy	AAGGGTCAACCATCCAAACCGTTCGTTGGCGTGTGAGCGCAGGTATTTAAACGCCGCCAGT	2340
2281	Db	AAGGGTCAACCATCCAAACCGTTCGTTGGCGTGTGAGCGCAGGTATTTAAACGCCGCCAGT	2340
2341	Qy	CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACCTATCTGTGACTGATGAAGGTCTG	2400
2341	Db	CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACCTATCTGTGACTGATGAAGGTCTG	2400
2401	Qy	GAAGCGGTTAATAAGACAAACCGCTGGGTGCCGTAGCGTGAAGTCTTACGAGGAAGAG	2460
2401	Db	GAAGCGGTTAATAAGACAAACCGCTGGGTGCCGTAGCGTGAAGTCTTACGAGGAAGAG	2460
2461	Qy	TTGCGGAAAGATCCACGCTATTGCGCGCACCATGGAACCGCCCAAGAGGTGAATCATG	2520
2461	Db	TTGCGGAAAGATCCACGCTATTGCGCGCACCATGGAACCGCCCAAGAGGTGAATCATG	2520
2521	Qy	CCGAACATCCGCGAGATGTCGCTTTTCTGGTATCCGCTGGCTACTCGGCTGATCAACGCC	2580
2521	Db	CCGAACATCCGCGAGATGTCGCTTTTCTGGTATCCGCTGGCTACTCGGCTGATCAACGCC	2580
2581	Qy	GCCAGCGGTGCTCAGACTGTGCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
2581	Db	GCCAGCGGTGCTCAGACTGTGCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG	2640
2641	Qy	AACAAACAAACAATACAATAAACAACAACCTCGGGATCGAGGGAAGATTTTCAGAATTC	2700
2641	Db	AACAAACAAACAATACAATAAACAACAACCTCGGGATCGAGGGAAGATTTTCAGAATTC	2700

RESULT 2

US-09-096-570-1
; Sequence 1, Application US/09096570
; Patent No. 6018020
; GENERAL INFORMATION:
; APPLICANT: Attwood, Michael R
; APPLICANT: Hurst, David N
; APPLICANT: Jones, Philip S
; APPLICANT: Kay, Paul B
; APPLICANT: Raynham, Tony M
; APPLICANT: Wilson, Francis X
; TITLE OF INVENTION: Amino Acid Derivatives
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: N.J.
; COUNTRY: U.S.A.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/096,570
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/971,036
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Kreisler, Lewis J
; REGISTRATION NUMBER: 38522

; LENGTH: 7475									
; TYPE: DNA									
; ORGANISM: Artificial Sequence									
; FEATURE:									
; OTHER INFORMATION: Description of Artificial Sequence: Vector and									
; OTHER INFORMATION: Gene Fragments									
US-09-265-617B-1									
Query Match 81.8%; Score 2700; DB 3; Length 7475;									
Best Local Similarity 100.0%; Pred. No. 0;									
Matches 2700; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
Qy	1	CGGACACCATCGAATGGTGCRAAAACCTTTCCGGGTATGSCATGATAGCGCCGGGAAGA	60						
Db	1	CGGACACCATCGAATGGTGCRAAAACCTTTCCGGGTATGSCATGATAGCGCCGGGAAGA	60						
Qy	61	GTCAATTCAGGGTGTGAATGTGAACACCAAGTAACGTATATACGATGTCGACAGTATGCCG	120						
Db	61	GTCAATTCAGGGTGTGAATGTGAACACCAAGTAACGTATATACGATGTCGACAGTATGCCG	120						
Qy	121	GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACACGAGCCAGCCACGTTTCTGCGAAAA	180						
Db	121	GTGTCCTTATCAGACCGTTTCCCGCGTGTGAACACGAGCCAGCCACGTTTCTGCGAAAA	180						
Qy	181	CGCGGAAAAAGTGAAGCGGCGATGGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC	240						
Db	181	CGCGGAAAAAGTGAAGCGGCGATGGCGGAGCTGAATTTACATTTCCCAACCGCGTGGCAC	240						
Qy	241	AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTTGCCACCTCCAGTCTGGGCCCTGC	300						
Db	241	AACAACTGGCGGCAAAACAGTCGTTGCTGATTTGGCGTTTGCCACCTCCAGTCTGGGCCCTGC	300						
Qy	301	ACGCGCGTGTGCAAAATTTGTGCGGCGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG	360						
Db	301	ACGCGCGTGTGCAAAATTTGTGCGGCGATTAATCTCGCGCCGATCAACTGGGTGCCAGCG	360						
Qy	361	TGTTGCTCGATGTGAGAACGAGCGGCTGAGAGCTGTAAAGCGGCGGTGCAATC	420						
Db	361	TGTTGCTCGATGTGAGAACGAGCGGCTGAGAGCTGTAAAGCGGCGGTGCAATC	420						
Qy	421	TTCTCGCGCAACCGCTCAGTGGCTGATCATTAATCTATCCCGTGGATGACACAGGATGCCA	480						
Db	421	TTCTCGCGCAACCGCTCAGTGGCTGATCATTAATCTATCCCGTGGATGACACAGGATGCCA	480						
Qy	481	TTGCTGTGGAAGCTGCCTGCACATAATTTCCGGCGTATTTCTTGATGTTCTCTGACCCAGA	540						
Db	481	TTGCTGTGGAAGCTGCCTGCACATAATTTCCGGCGTATTTCTTGATGTTCTCTGACCCAGA	540						
Qy	541	CACCCATCAACAGTATTTATTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600						
Db	541	CACCCATCAACAGTATTTATTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC	600						
Qy	601	TGGTCGCATTTGGTCCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTCTGTCTCGG	660						
Db	601	TGGTCGCATTTGGTCCACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTCTGTCTCGG	660						
Qy	661	CGCGTCTGCGTCTGGCTGSCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
Db	661	CGCGTCTGCGTCTGGCTGSCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
Qy	721	CGGAACGGGAAGCGACGTGAGTGCATGTCGGTTTTTCAAAACCATGCAAAATGCTGA	780						
Db	721	CGGAACGGGAAGCGACGTGAGTGCATGTCGGTTTTTCAAAACCATGCAAAATGCTGA	780						
Qy	781	ATGAGGGCATCGTCCCACTGCGATGCTGGTTGCGCAACGATCAGATGGCGCTGGCGCAA	840						
Db	781	ATGAGGGCATCGTCCCACTGCGATGCTGGTTGCGCAACGATCAGATGGCGCTGGCGCAA	840						
Qy	841	TGCGGCGCATTAACCGAGTCCGGGCTCGCGTTTGGTGGGATATCTCGGTAGTGGGATAGC	900						
Db	841	TGCGGCGCATTAACCGAGTCCGGGCTCGCGTTTGGTGGGATATCTCGGTAGTGGGATAGC	900						
Qy	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCATCAAAAGGATTTTC	960						

Db	901	ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCATCAAAAGGATTTTC	960						
Qy	961	GCCTGCTGGGCAAAACAGCGTGAACCGCTTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020						
Db	961	GCCTGCTGGGCAAAACAGCGTGAACCGCTTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020						
Qy	1021	AGGGCAATCAGCTGTGCCCCGTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCAATA	1080						
Db	1021	AGGGCAATCAGCTGTGCCCCGTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCAATA	1080						
Qy	1081	CGCAAAACCGCTCTCTCCCGCGCTTGGCCGATTCATTAAATGAGCTGGCAACAGAGTTT	1140						
Db	1081	CGCAAAACCGCTCTCTCCCGCGCTTGGCCGATTCATTAAATGAGCTGGCAACAGAGTTT	1140						
Qy	1141	CCGCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG	1200						
Db	1141	CCGCACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGCTCACTCATTTAG	1200						
Qy	1201	GCACAAATTCATGTTTGACAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGG	1260						
Db	1201	GCACAAATTCATGTTTGACAGCTTATCATCGACTGCAAGTGCACCAATGCTTCTGGG	1260						
Qy	1261	TCAGGACGCAATCGGAAGCTGTGGTATGCGCTGTCAGGTCTGTAATCACTGCATAATTG	1320						
Db	1261	TCAGGACGCAATCGGAAGCTGTGGTATGCGCTGTCAGGTCTGTAATCACTGCATAATTG	1320						
Qy	1321	TGTGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTGGCGGACATCAATACGGTT	1380						
Db	1321	TGTGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTGGCGGACATCAATACGGTT	1380						
Qy	1381	CTGGCAATATTTGAAATGAGCTGTGCAATTAATCATCGGCTCGTATATGTTGGA	1440						
Db	1381	CTGGCAATATTTGAAATGAGCTGTGCAATTAATCATCGGCTCGTATATGTTGGA	1440						
Qy	1441	ATTGTAGCGGATTAACAATTTTCAACAGAAACAGCCAGTCCGTTTGGTGTTCACGA	1500						
Db	1441	ATTGTAGCGGATTAACAATTTTCAACAGAAACAGCCAGTCCGTTTGGTGTTCACGA	1500						
Qy	1501	GCACCTTCAACCAAGGACCATAGATTATGAATTAAGAAAGTGAATCTGGTAATCTGG	1560						
Db	1501	GCACCTTCAACCAAGGACCATAGATTATGAATTAAGAAAGTGAATCTGGTAATCTGG	1560						
Qy	1561	ATTTAAGCGGATTAAGGCTATAACCGTCTCGTGAAGTCGGTGAAGAAATTCAGAAAAGT	1620						
Db	1561	ATTTAAGCGGATTAAGGCTATAACCGTCTCGTGAAGTCGGTGAAGAAATTCAGAAAAGT	1620						
Qy	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGATT	1680						
Db	1621	ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTCACAGATT	1680						
Qy	1681	CGCGCAACTGGCGATGGCCCTGACATTAATCTTTGGGCACACGACCGCTTTGGTGGCTAC	1740						
Db	1681	CGCGCAACTGGCGATGGCCCTGACATTAATCTTTGGGCACACGACCGCTTTGGTGGCTAC	1740						
Qy	1741	GCTCAATTCGCGCTGTGGCTGAAATCAACCCGGCAAAAGCGTTCCAGACAAGCTGTAT	1800						
Db	1741	GCTCAATTCGCGCTGTGGCTGAAATCAACCCGGCAAAAGCGTTCCAGACAAGCTGTAT	1800						
Qy	1801	CCGTTTACCTGGATGCGGTACGTTTACAAACGCAAGCTGATTGCTTACCGATCGCTGT	1860						
Db	1801	CCGTTTACCTGGATGCGGTACGTTTACAAACGCAAGCTGATTGCTTACCGATCGCTGT	1860						
Qy	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA	1920						
Db	1861	GAAGCGTTATCGCTGATTTTAAACAAAGATCTGCTGCCGAAACCCGCCAAAAACCTGGGAA	1920						
Qy	1921	GAGATCCCGCGCTGGATAAGAACTGAAGAGAAAGTGAAGCGCGCTGATGTTCAAC	1980						
Db	1921	GAGATCCCGCGCTGGATAAGAACTGAAGAGAAAGTGAAGCGCGCTGATGTTCAAC	1980						
Qy	1981	CTGCAAGAACCGTACTTACCTGGCGCTGATTGCTGTCGACGGGGTTATCGGTTCAAG	2040						

1981	CTGCAAGAACGGTACTTCACTGCGCGGTGATTGCTGCTGACGGGGGTATTGCGTTCAAG	2041
Qy	TATGAAAAACGGCAAGTAGTCAGCATTTAAAGACGTTGGCGCTGGATACGCTGGCGCGAAAGCG	2100
Db	TATGAAAAACGGCAAGTAGTCAGCATTTAAAGACGTTGGCGCTGGATACGCTGGCGCGAAAGCG	2160
Qy	GGTCTGACCTTCTGGTTGACCTGATTTAAAAAACAACACATGATGACAGACACCGATTAC	2220
Db	GGTCTGACCTTCTGGTTGACCTGATTTAAAAAACAACACATGATGACAGACACCGATTAC	2280
Qy	TCCATCGCAAGAGCTGCCTTTTATAAAGGCGAAACACGCGATGACCATCAACGCGCCGTGG	2340
Db	TCCATCGCAAGAGCTGCCTTTTATAAAGGCGAAACACGCGATGACCATCAACGCGCCGTGG	2400
Qy	GCATGTCCAACATTCGACACACAGCAAAAGTGAATTTATGGTGTAAACGGTACTGCCGACCTTC	2460
Db	GCATGTCCAACATTCGACACACAGCAAAAGTGAATTTATGGTGTAAACGGTACTGCCGACCTTC	2520
Qy	AAGGGTCAACCATCCAAACCGTTTCGTTGGCGTCTGAGCGCAGGTATTTAAGCGCGCCAGT	2580
Db	AAGGGTCAACCATCCAAACCGTTTCGTTGGCGTCTGAGCGCAGGTATTTAAGCGCGCCAGT	2640
Qy	CCGAAACAAGAGCTGGCAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG	2700
Db	CCGAAACAAGAGCTGGCAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG	2760
Qy	GAAGCGGTTAATAAGACAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAGAG	2820
Db	GAAGCGGTTAATAAGACAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAGAG	2880
Qy	TTGCGGAAAGATCCACGTATTTCGCGCCACCATGGAAAAACGCCAGAAAGGTGAATCATG	2940
Db	TTGCGGAAAGATCCACGTATTTCGCGCCACCATGGAAAAACGCCAGAAAGGTGAATCATG	3000
Qy	CCGAAACATCCCGCAGATGTCGGCTTTCGTGTATCCGTGGGTACTGGGTGATCAAGCC	3060
Db	CCGAAACATCCCGCAGATGTCGGCTTTCGTGTATCCGTGGGTACTGGGTGATCAAGCC	3120
Qy	GCCAGCGGTGCTCAGACTGTCGATGAAGCCCTGAAAGACCGCGACAGACTAATTCGAGCTCG	3180
Db	GCCAGCGGTGCTCAGACTGTCGATGAAGCCCTGAAAGACCGCGACAGACTAATTCGAGCTCG	3240
Qy	AACCAACAACAATTAACAATAACAACCTCGGATCGAGGGAAGGATTTTCAGAAATTC	3300
Db	AACCAACAACAATTAACAATAACAACCTCGGATCGAGGGAAGGATTTTCAGAAATTC	3360

RESULT 4

US-09-027-169-3/C
Sequence 3, Application US/09027169
Patent No. 6420524
GENERAL INFORMATION:
APPLICANT: CRAIG, NANCY L
TITLE OF INVENTION: GAIN OF FUNCTION MUTATIONS IN
TITLE OF INVENTION: ATP-DEPENDENT TRANSCRIPTION PROTEINS
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Anne Brown (Alston & Bird, LLP)
STREET: 3605 Glenwood Ave.
CITY: Raleigh
STATE: NC
COUNTRY: USA
ZIP: 27608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/027,169
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

|||||
784 GAGGCATCGTTCACATGCGATGCTGGTGGCAACGATCAGATGGCGTGGCGCAATG 725
Qy CGCGCATTTACCGAGTCGGCGTGGCGGTGGTGGCGGATATCTCGGTAGTGGATACGAC 902
Db CGCGCATTTACCGAGTCGGCGTGGCGGTGGTGGCGGATATCTCGGTAGTGGATACGAC 665
Qy GATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAACAGGATTTTCG 962
Db GATACCGAAGACAGCTCATGTTATATCCCGCGTCAACCAACCATCAACAGGATTTTCG 605
Qy CTGCTGGGGCAAAACAGCGTGGACCGTTCCTGCAACTCTCTCAGGGCCAGCGGTGAAG 1022
Db CTGCTGGGGCAAAACAGCGTGGACCGTTCCTGCAACTCTCTCAGGGCCAGCGGTGAAG 545
Qy GGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAGAAACACACCTCGCGCCCAATACG 1082
Db GGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAGAAACACACCTCGCGCCCAATACG 485
Qy CAAACCGCTCTCCCGCGGTGGCGGATTCATTAATGCGAGCTGGCAGCAGGTTTCC 1142
Db CAAACCGCTCTCCCGCGGTGGCGGATTCATTAATGCGAGCTGGCAGCAGGTTTCC 425
Qy CGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTGAGTTAGCTCACTATTAGGC 1202
Db CGACTGGAAGCGGCGAGTGAAGCGCAACGCAATTAATGTGAGTTAGC-----GC 376
Qy ACAATTCTCATGTTTGACAGCTTATCATCGACTGACCGGTGACCAATGCTTCTGGCGTC 1262
Db GAAATTGATCTGGTTGACAGCTTATCATCGACTGACCGGTGACCAATGCTTCTGGCGTC 316
Qy AGGAGCCATCGGAAGCTGTGGTATGGCTGTGCGAGTGGTAAATCACTGCATAATTCGTG 1322
Db AGGAGCCATCGGAAGCTGTGGTATGGCTGTGCGAGTGGTAAATCACTGCATAATTCGTG 256
Qy TCGCTCAAGGGCGCACTCCGCTCTGGATAATGTTTTTGGCGCGACATCAACGGTTCT 1382
Db TCGCTCAAGGGCGCACTCCGCTCTGGATAATGTTTTTGGCGCGACATCAACGGTTCT 196
Qy GGCAATATTCGAAATGAGCTGTTGACAAATTAATCAT- CGGCTCGTATAATGTGGAA 1441
Db GGCAATATTCGAAATGAGCTGTTGACAAATTAATCATCAATCATCCGGCTCGTATAATGTGGAA 136
Qy TTGTGAGCGGATAACAATTTACACAGGAAACAGCC 1477
Db TTGTGAGCGGATAACAATTTACACAGGAAACAGAC 100

RESULT 5
US-09-640-882-2
; Sequence 2, Application US/09640882
; Patent No. 6720142
; GENERAL INFORMATION:
; APPLICANT: Hall, Barry G.
; TITLE OF INVENTION: METHOD OF DETERMINING EVOLUTIONARY POTENTIAL OF MUTANT
; TITLE OF INVENTION: RESISTANCE GENES AND USE THEREOF TO SCREEN FOR DRUG
; TITLE OF INVENTION: EFFICACY
; FILE REFERENCE: 176/60851
; CURRENT APPLICATION NUMBER: US/09/640,882
; PRIOR FILING DATE: 2000-08-18
; PRIOR APPLICATION NUMBER: 60/149,813
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 5201
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: plasmid
; OTHER INFORMATION: PACSE2
US-09-640-882-2

Query Match 41.3%; Score 1362; DB 4; Length 5201;
Best Local Similarity 99.1%; Pred. No. 0;
Matches 1380; Conservative 0; Mismatches 10; Indels 2; Gaps 1;
Qy 3 GACACCATCGAATGGTGCAAAACCTTTTCGCGGTATGGCATATAGCGCCGCCGAAAGAGT 62
Db 866 GACACCATCGAATGGCGCAAAACCTTTTCGCGGTATGGCATATAGCGCCGCCGAAAGAGT 925
Qy 63 CAATTCAGGGTGGTGAATGTGAAACCAAGTAACTGATATACGATGTCGAGAGTATGCCGGT 122
Db 926 CAATTCAGGGTGGTGAATGTGAAACCAAGTAACTGATATACGATGTCGAGAGTATGCCGGT 985
Qy 123 GTCTCTTATCAGACGGTTCCTCGCGTGGTGAACAGGCGCCAGCTTTCTGCGGAAACG 182
Db 986 GTCTCTTATCAGACGGTTCCTCGCGTGGTGAACAGGCGCCAGCTTTCTGCGGAAACG 1045
Qy 183 CGGGAAGGTTGGAAGCGGCGATGGCGGAGCTGAAATTAATTCATTTCCCAACCGGTGCAAA 242
Db 1046 CGGGAAGGTTGGAAGCGGCGATGGCGGAGCTGAAATTAATTCATTTCCCAACCGGTGCAAA 1105
Qy 243 CAATTCAGGGCGCAACAGTCTGATTCGCGGTGGCCACTCCAGTCTGGCCCTGCAC 302
Db 1106 CAATTCAGGGCGCAACAGTCTGATTCGCGGTGGCCACTCCAGTCTGGCCCTGCAC 1165
Qy 303 CGCGCTCGCAAAATGTCGCGCGGATTAATTCGCGCGGATCACTCGCGCGGATCACTGGGTGCCAGCGTG 362
Db 1166 CGCGCTCGCAAAATGTCGCGCGGATTAATTCGCGCGGATCACTCGCGCGGATCACTGGGTGCCAGCGTG 1225
Qy 363 GTGGTGTGATGTTAGAACGCGGCGTGAAGCGCTGTAAGCGGCGGTGCACAACTTT 422
Db 1226 GTGGTGTGATGTTAGAACGCGGCGTGAAGCGCTGTAAGCGGCGGTGCACAACTTT 1285
Qy 423 CTCGCGCAACCGTCACTGGGCTGATCAATTAATTCGCGCTGGATGACAGGATGCCATT 482
Db 1286 CTCGCGCAACCGTCACTGGGCTGATCAATTAATTCGCGCTGGATGACAGGATGCCATT 1345
Qy 483 GCTGTGGAGCTGCTGCACTAAATGTTCCGCGGTATTTCTTGATGTCTGTACAGACA 542
Db 1346 GCTGTGGAGCTGCTGCACTAAATGTTCCGCGGTATTTCTTGATGTCTGTACAGACA 1405
Qy 543 CCCATCAACAGTATTTCTCCATGAAGACGTACGCGCTGGCGGTGGAGCATCTG 602
Db 1406 CCCATCAACAGTATTTCTCCATGAAGACGTACGCGCTGGCGGTGGAGCATCTG 1465
Qy 603 CTCGCAATTTGGTCAACAGCAAAATCGCGTGTGTAGCGGCGCCATTAAGTTCTGTCTCGGG 662
Db 1466 CTCGCAATTTGGTCAACAGCAAAATCGCGTGTGTAGCGGCGCCATTAAGTTCTGTCTCGGG 1525
Qy 663 CGTCTGCGTCTGGCTGGCGTGAATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
Db 1526 CGTCTGCGTCTGGCTGGCGTGAATAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 1585
Qy 723 GAACGGGAAGCGACTGGAGTGCATGTCGCGTTCACAAACCATCAAAATTCGCAATCTGAAT 782
Db 1586 GAACGGGAAGCGACTGGAGTGCATGTCGCGTTCACAAACCATCAAAATTCGCAATCTGAAT 1645
Qy 783 GAGGCGATCTGTTCCCATCTGCGATGCTGTTGTCACCAACGATCAGATGGCGCTGGCGCAATG 842
Db 1646 GAGGCGATCTGTTCCCATCTGCGATGCTGTTGTCACCAACGATCAGATGGCGCTGGCGCAATG 1705
Qy 843 CGCGCAATTCAGAGTCCGGCTGCGCGTGGTGGGATATCTCGGTAGTGGATACGAC 902
Db 1706 CGCGCAATTCAGAGTCCGGCTGCGCGTGGTGGGATATCTCGGTAGTGGATACGAC 1765
Qy 903 GATACCGAAGACGCTCATGTTATATCCCGCGTGAACCAACCATCAAAACAGATTTTCG 962
Db 1766 GATACCGAAGACGCTCATGTTATATCCCGCGTGAACCAACCATCAAAACAGATTTTCG 1825
Qy 963 CTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCGAGCGGTGAAG 1022
Db 1826 CTGCTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCGAGCGGTGAAG 1885
Qy 1023 GGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAGAAACCAACCTCGCGCCCAATACG 1082

Db 1886 GGCATATCAGCTGTTGCGGCTCTCACTGGTGAAGAAAGAAAACCACTCGCGCCCAATAGC 1945
 Qy 1083 CAAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGCGAGCTGGCGAGCAGGTTTCC 1142
 Db 1946 CAAACCGCTCTCCCGCGGCTTGGCGGATTCATTAATGCGAGCTGGCGAGCAGGTTTCC 2005
 Qy 1143 CCACTGGAAGCGGCGAGTGAAGCGCAAGCAATTAATGTAGTTAGCTCACTCATTAGGC 1202
 Db 2006 CCACTGGAAGCGGCGAGTGAAGCGCAAGCAATTAATGTAGTTAG--CGGAATTGATC 2063
 Qy 1203 ACAATTCATGTTTGACAGCTTATCATCGACTGACGCTGACGAGTGAACCAATGCTTGGCGTC 1262
 Db 2064 TGAATTCATGTTTGACAGCTTATCATCGACTGACGCTGACGAGTGAACCAATGCTTCTGGCGTC 2123
 Qy 1263 AGGAGCCATCGGAAGCTGTGTATGCTGTGTCAGGTGTAATCACTGCATGAATTCGTG 1322
 Db 2124 AGGAGCCATCGGAAGCTGTGTATGCTGTGTCAGGTGTAATCACTGCATGAATTCGTG 2183
 Qy 1323 TCGCTCAAGGCGCACTCCCGTCTGGATTAATGTTTTCGCGCGACATCAATAGCGTTCT 1382
 Db 2184 TCGCTCAAGGCGCACTCCCGTCTGGATTAATGTTTTCGCGCGACATCAATAGCGTTCT 2243
 Qy 1383 GGCATAATTTCT 1394
 Db 2244 GGCATAATTTCT 2255

RESULT 6

US-09-640-882-3
 ; Sequence 3, Application US/09640882
 ; Patent No. 6720142
 ; GENERAL INFORMATION:
 ; APPLICANT: Hall, Barry G.
 ; TITLE OF INVENTION: METHOD OF DETERMINING EVOLUTIONARY POTENTIAL OF MUTANT
 ; TITLE OF INVENTION: RESISTANCE GENES AND USE THEREOF TO SCREEN FOR DRUG
 ; TITLE OF INVENTION: EFFICACY
 ; FILE REFERENCE: 176/60851
 ; CURRENT APPLICATION NUMBER: US/09/640,882
 ; CURRENT FILING DATE: 2000-08-18
 ; PRIOR APPLICATION NUMBER: 60/149,813
 ; PRIOR FILING DATE: 1999-08-19
 ; NUMBER OF SEQ ID NOS: 17
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 5201
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: plasmid
 ; OTHER INFORMATION: PACSE3
 US-09-640-882-3

Query Match 41.3%; Score 1362; DB 4; Length 5201;
 Best Local Similarity 99.1%; Pred. No. 0;
 Matches 1380; Conservative 0; Mismatches 10; Indels 2; Gaps 1;
 Qy 3 GACACCATCGAATGGTGCAGAAACCTTTTCGCGTATGGCATGATAGCGCCCGGAAGAGAGT 62
 Db 866 GACACCATCGAATGGCGCAAAACCTTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGAGT 925
 Qy 63 CAATTTCAGGTGGTGAATGTGAACCAAGTAACGTTATACCATGTCGAGAGTATGCCGGT 122
 Db 926 CAATTTCAGGTGGTGAATGTGAACCAAGTAACGTTATACCATGTCGAGAGTATGCCGGT 985
 Qy 123 GTCTCTTATCAGACCGTTTCCCGGTGGTGAACCAAGCGCCAGCCACGTTTCTGCGAAACG 182
 Db 986 GTCTCTTATCAGACCGTTTCCCGGTGGTGAACCAAGCGCCAGCCACGTTTCTGCGAAACG 1045
 Qy 183 CGGGAAGTGAAGCGGCGATGGCGAGCTGAATTAATTCACCAACCGCTGGCACA 242
 Db 1046 CGGGAAGTGAAGCGGCGATGGCGAGCTGAATTAATTCACCAACCGCTGGCACA 1105

Qy 243 CAACCTGGCGGCAACAGTCTGTTGATTTGGCGTTGGACCTCCAGTCTGGCCCTCGAC 302
 Db 1106 CAACCTGGCGGCAACAGTCTGTTGATTTGGCGTTGGACCTCCAGTCTGGCCCTCGAC 1165
 Qy 303 GCGCGCTGCAAAATTTGTCGCGGCGATTTAAATCTCGCCCGGATCAATCTGGGTGCCAGCGTG 362
 Db 1166 GCGCGCTGCAAAATTTGTCGCGGCGATTTAAATCTCGCCCGGATCAATCTGGGTGCCAGCGTG 1225
 Qy 363 GTGCTGCGATGTGTAAGAACGAAGCGGCTGGAAGCCTGTAAAGCGGCGGTGCACAATCTT 422
 Db 1226 GTGCTGCGATGTGTAAGAACGAAGCGGCTGGAAGCCTGTAAAGCGGCGGTGCACAATCTT 1285
 Qy 423 CTGCGCAACGCGTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACAGGATGCGCAT 482
 Db 1286 CTGCGCAACGCGTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACAGGATGCGCAT 1345
 Qy 483 GCTGTGGAAGCTGCTGCACTAATGTTCCGCGCTTATTTCTTGTATGTTCTCTGACCAAGACA 542
 Db 1346 GCTGTGGAAGCTGCTGCACTAATGTTCCGCGCTTATTTCTTGTATGTTCTCTGACCAAGACA 1405
 Qy 543 CCCATCAACAGTATTTTCTCCATGAAGACGCTAGCGCATGGGCGTGGAGCATCTG 602
 Db 1406 CCCATCAACAGTATTTTCTCCATGAAGACGCTAGCGCATGGGCGTGGAGCATCTG 1465
 Qy 603 GTGCGATTGGGTCAACAGCAAAATCGCGCTGTAGCGGCGCATTTAAGTTCTGTTCTCGCG 662
 Db 1466 GTGCGATTGGGTCAACAGCAAAATCGCGCTGTAGCGGCGCATTTAAGTTCTGTTCTCGCG 1525
 Qy 663 CGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGCATAGCG 722
 Db 1526 CGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGCATAGCG 1585
 Qy 723 GAACGGGAAGCGCACTGGATGCGCATGTCGGTTCCTCAACAAACCATGCAATGCTGAAT 782
 Db 1586 GAACGGGAAGCGCACTGGATGCGCATGTCGGTTCCTCAACAAACCATGCAATGCTGAAT 1645
 Qy 783 GAGGCGATCGTTCCTCCACTGCGATGCTGTTGTCGCAACGATCAGATGCGCTGGCGCGCAATG 842
 Db 1646 GAGGCGATCGTTCCTCCACTGCGATGCTGTTGTCGCAACGATCAGATGCGCTGGCGCGCAATG 1705
 Qy 843 GCGGCCATTACCGAGTCCGGGCTGGCGGTTGGTGCGGATATCTCGGTAGTGGGATACGAC 902
 Db 1706 GCGGCCATTACCGAGTCCGGGCTGGCGGTTGGTGCGGATATCTCGGTAGTGGGATACGAC 1765
 Qy 903 GATACCGAAGACAGCTCATGTTTATATCCCGCGGTAAACCAACCATCAACAGGATTTTCG 962
 Db 1766 GATACCGAAGACAGCTCATGTTTATATCCCGCGGTAAACCAACCATCAACAGGATTTTCG 1825
 Qy 963 CTGCTGGGCAACACGAGCTGGAACCGCTTGTGTCGCACTCTCTCAGGGCCAGCGGTGAAG 1022
 Db 1826 CTGCTGGGCAACACGAGCTGGAACCGCTTGTGTCGCACTCTCTCAGGGCCAGCGGTGAAG 1885
 Qy 1023 GGCATACAGTGTGGCGCTCTCACTGGTGAAGAAAGAAAACCAACCTGGCGCCCAATACG 1082
 Db 1886 GGCATACAGTGTGGCGCTCTCACTGGTGAAGAAAGAAAACCAACCTGGCGCCCAATACG 1945
 Qy 1083 CAAACCGCTCTCCCGCGCGTTGGCGGATTCATTAATCAGCTGGCAGCAGAGGTTTCC 1142
 Db 1946 CAAACCGCTCTCCCGCGCGTTGGCGGATTCATTAATCAGCTGGCAGCAGAGGTTTCC 2005
 Qy 1143 CGACTGGAAGCGGCGAGTGAAGCGCAACCGCAATTAATGTAGTTAGCTCACTCATTAGGC 1202
 Db 2006 CGACTGGAAGCGGCGAGTGAAGCGCAACCGCAATTAATGTAGTTAG--CGCAATTTGATC 2063
 Qy 1203 ACAATTCATGTTTGACAGCTTATCATCGACTGACGCTGACCAATGCTTCTGGCGTC 1262
 Db 2064 TGAATTCATGTTTGACAGCTTATCATCGACTGACGCTGACCAATGCTTCTGGCGTC 2123
 Qy 1263 AGGAGCCATCGGAAGCTGTGTTATGGCTGTGAGGTGTCAGGTGCTAAATCACTGCATAATTCGTG 1322
 Db 2124 AGGAGCCATCGGAAGCTGTGTTATGGCTGTGAGGTGTCAGGTGCTAAATCACTGCATAATTCGTG 2183
 Qy 1323 TCGCTCAAGCGCAGCTCCCGTCTCGATAATGTTTTCGCGCGACATCATTAACGGTTCT 1382

Db	2184	TCGCTCAAGGCGCACTCCCGTTCGGATAATGTTTTTGGCCCGACATCATACCGTTCT	2243
Qy	1383	GGCAATATTCT 1394	
Db	2244	GGCAATATTCT 2255	
RESULT 7			
US-08-778-717-5			
; Sequence 5, Application US/08778717			
; Patent No. 6602689			
; GENERAL INFORMATION:			
; APPLICANT: UENO, EIICHI			
; APPLICANT: NOBUYUKI, FUJII			
; APPLICANT: OKADA, MASAHISA			
; TITLE OF INVENTION: FUSED DNA SEQUENCE, FUSED PROTEIN			
; TITLE OF INVENTION: EXPRESSED FROM SAID FUSED DNA SEQUENCE AND METHOD FOR			
; TITLE OF INVENTION: EXPRESSING SAID DNA SEQUENCE			
; NUMBER OF SEQUENCES: 21			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,			
; ADDRESS: P.C.			
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400			
; CITY: ARLINGTON			
; STATE: VA			
; COUNTRY: USA			
; ZIP: 22202			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: Floppy disk			
; COMPUTER: IBM PC compatible			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: PatentIn Release #1.0, Version #1.30			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/08/778,717			
; FILING DATE: 12-DEC-1996			
; CLASSIFICATION: 536			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: JP 352225/1995			
; FILING DATE: 28-DEC-1995			
; ATTORNEY/AGENT INFORMATION:			
; NAME: OBLON, NORMAN F.			
; REGISTRATION NUMBER: 24,618			
; REFERENCE/DOCKET NUMBER: 2084-031-0			
; TELECOMMUNICATION INFORMATION:			
; TELEPHONE: 703-413-3000			
; TELEFAX: 703-413-2220			
; INFORMATION FOR SEQ ID NO: 5:			
; SEQUENCE CHARACTERISTICS:			
; LENGTH: 4557 base pairs			
; TYPE: nucleic acid			
; STRANDEDNESS: double			
; TOPOLOGY: circular			
; MOLECULE TYPE: other nucleic acid			
; ORIGINAL SOURCE:			
; ORGANISM: E. COLI			
; STRAIN: BL21 (DE3)			
; PUBLICATION INFORMATION:			
; AUTHORS: NOBUYUKI FUJII ET AL,			
; TITLE: FUSED DNA SEQUENCE, FUSED PROTEIN EXPRESSED			
; TITLE: FROM SAID FUSED DNA SEQUENCE AND METHOD OF			
; TITLE: EXPRESSING SAID FUSED PROTEIN			
; RELEVANT RESIDUES IN SEQ ID NO: 5: FROM 1 TO 4557			
US-08-778-717-5			
Query Match 36.5%; Score 1204; DB 4; Length 4557;			
Best Local Similarity 100.0%; Pred. No. 0;			
Matches 1204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	CCGACACATCGAATGGTGCNAAAACCTTTCCGGGTATGGCATGATAGGCCCGGAAGAGA	60
Db	2563	CCGACACATCGAATGGTGCNAAAACCTTTCCGGGTATGGCATGATAGGCCCGGAAGAGA	2622

Db 3703 CCCGACTGGAAAGCGGCGAGTGGCGCAACGCAATTAATGTAGTACTCACTATTAG 3762
QY 1201 GCAC 1204
Db 3763 GCAC 3766

RESULT 8
US-08-148-675A-2
; Sequence 2, Application US/08148675A
; Patent No. 5506121
; GENERAL INFORMATION:
; APPLICANT: Skerita, Arne; Schmidt, Thomas
; TITLE OF INVENTION: FUSION PEPTIDES WITH BINDING ACTIVITY FOR
; TITLE OF INVENTION: STREPTAVIDIN
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/148.675A
; FILING DATE: 3-NOV-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 42 37 113.9
; FILING DATE: 03-NOV-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Isal, Christine H.
; REGISTRATION NUMBER: 34,266
; REFERENCE/DOCKET NUMBER: HUBR 1041
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3832 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-148-675A-2

Query Match 36.4%; Score 1200.8; DB 1; Length 3832;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1202; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGTTGCAAAACCTTCCCGGTATGGCATGATAGCGCCGGAAGAGA 60
Db 3 CCGACACCATCGAATGCGCAAAACCTTTCGCGGTATGGCATGATAGCGCCGGAAGAGA 62
QY 61 GTCAATTGAGGGTGGTGAATGTGAACCAAGTAACTTATACGATGTCGACAGATATGCCG 120
Db 63 GTCAATTGAGGGTGGTGAATGTGAACCAAGTAACTTATACGATGTCGACAGATATGCCG 122
QY 121 GTGTCTCTTATCAGACCGTTTCCCGGTGTTGAACCAAGTAACTTATCCTCCAAACCGGTGGCAC 180
Db 123 GTGTCTCTTATCAGACCGTTTCCCGGTGTTGAACCAAGTAACTTATCCTCCAAACCGGTGGCAC 182
QY 181 CGCGGGAAGAGTGGACCGCGATGGCGAGCTGAATTACATCCCAACCGGTGGCAC 240
Db 183 CGCGGGAAGAGTGGACCGCGATGGCGAGCTGAATTACATCCCAACCGGTGGCAC 242
QY 241 AACAACTGGCGGCAACAGTCGTTGCTGATTTGGCGTTGCCACTCCAGTCTGGCCCTGC 300

Db 243 AACAACTGGCGGCAACAGTCGTTGCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 302
QY 301 ACGGCGCGTCCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 360
Db 303 ACGGCGCGTCCGCAAAATTTGTCGGCGGATTAATCTCGCGCGGATCAACTGGGTGCCAGCG 362
QY 361 TGGTGGTGTGATGTTAGAACGAAAGCGGCTGCGAAGCTCTGTAAGCGCGGTGCACAATC 420
Db 363 TGGTGGTGTGATGTTAGAACGAAAGCGGCTGCGAAGCTCTGTAAGCGCGGTGCACAATC 422
QY 421 TTCTCGGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGATGACGAGATGCCA 480
Db 423 TTCTCGGCAACCGCTCAGTGGGCTGATCAATTAATCTCCGCTGATGACGAGATGCCA 482
QY 481 TTGCTGTGGAGCTGCGCTGCACTAATTTTCGGGGTTATTTCTTGATGTCTCTGACCAGA 540
Db 483 TTGCTGTGGAGCTGCGCTGCACTAATTTTCGGGGTTATTTCTTGATGTCTCTGACCAGA 542
QY 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Db 543 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 602
QY 601 TGGTGGCATTTGGTCCACGACAAATCGCGCTTTAGCGGGCCCAATTAAGTTCTCTCGG 660
Db 603 TGGTGGCATTTGGTCCACGAAATCGCGCTTTAGCGGGCCCAATTAAGTTCTCTCGG 662
QY 661 CGGCTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 663 CGGCTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 722
QY 721 CGGAACGGGAAGCGACTGGAGTGCCATGTCGCGTTTTCACAAACCATGCAATGCTGA 780
Db 723 CGGAACGGGAAGCGACTGGAGTGCCATGTCGCGTTTTCACAAACCATGCAATGCTGA 782
QY 781 ATGAGGCATCTGTTCCACATCGCGATGCTGTGCAACAGATAGATGGCGCTGGCGCAA 840
Db 783 ATGAGGCATCTGTTCCACATCGCGATGCTGTGCAACAGATAGATGGCGCAA 842
QY 841 TGGCGGCATTACCGAGTCCGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 843 TGGCGGCATTACCGAGTCCGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATACG 902
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACACCATCAAAACAGGATTTTC 960
Db 903 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACACCATCAAAACAGGATTTTC 962
QY 961 GCCTGTGGGGCAACCGAGCGTGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 963 GCCTGTGGGGCAACCGAGCGTGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1022
QY 1021 AGGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAACACCCCTGGCGCCCAATA 1080
Db 1023 AGGGCAATCAGCTGTTGCCGCTCTCACTGGTGAAGAAACACCCCTGGCGCCCAATA 1082
QY 1081 CGAAACCGCCTCTCCCGCGCTTGGCGGATTCATTAAATGCAAGCTGGCACGACAGGTTT 1140
Db 1083 CGAAACCGCCTCTCCCGCGCTTGGCGGATTCATTAAATGCAAGCTGGCACGACAGGTTT 1142
QY 1141 CCGGACTGGAAAGCGGCGAGTGACCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
Db 1143 CCGGACTGGAAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1202
QY 1201 GCAC 1204
Db 1203 GCAC 1206

RESULT 9
US-08-487-283A-18/c
; Sequence 18, Application US/08487283A
; Patent No. 6355245
; GENERAL INFORMATION:
; APPLICANT: Evans, Mark J.

APPLICANT: Matis, Louis A.
APPLICANT: Mueller, Eileen Elliott
APPLICANT: Nye, Steven H.
APPLICANT: Rollins, Scott
APPLICANT: Rother, Russell P.
APPLICANT: Springhorn, Jeremy P.
APPLICANT: Squinto, Stephen P.
APPLICANT: Thomas, Thomas C.
APPLICANT: Wilkins, James A.
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE TREATMENT
OF INFLAMMATORY DISEASES
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seth A. Fidel
STREET: 25 Science Park (Alexion)
CITY: New Haven
STATE: Connecticut
COUNTRY: USA
ZIP: 06511
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.4Mb storage
COMPUTER: Macintosh Cetrus 610
OPERATING SYSTEM: System 7
SOFTWARE: WordPerfect 3.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,283A
FILING DATE: June 7, 1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/236,208
FILING DATE: 02-MAY-1994
ATTORNEY/AGENT INFORMATION:
NAME: Seth A. Fidel.
REGISTRATION NUMBER: 38,449
REFERENCE/DOCKET NUMBER: ALX-152.1 CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (203)776-1790
TELEFAX: (203)772-3655
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 5248 base pairs
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Circular
MOLECULE TYPE: Other nucleic acid
DESCRIPTION: PET Trc SOS/NI
DESCRIPTION: prokaryotic expression vector
US-08-487-283A-18

Query Match 36.3%; Score 1199.2; DB 3; Length 5248;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTCGGGGTATGCGATAGCGCCGGAAGAGA 60
DB 4810 CGGACACCATCGAATGGCGCAAAACCTTTCGGGGTATGCGATAGCGCCGGAAGAGA 4751
QY 61 GTCATTCAGGGTGTGATGTGAACAGTAAACCGTTATACGATCTCGCAGAGTATGCG 120
DB 4750 GTCATTCAGGGTGTGATGTGAACAGTAAACCGTTATACGATCTCGCAGAGTATGCG 4691
QY 121 GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACAGCGCCAGCCACGTTTCTCGAAAA 180
DB 4690 GTGTCCTTATCAGACCGTTTCCCGGTGGTGAACAGCGCCAGCCGTTTCTCGAAAA 4631
QY 181 CGCGGAAAAAGTGAAGCGGATGGCGGAGCTGAATTACATTCGCAACCGGTGGCAC 240
DB 4630 CGCGGAAAAAGTGAAGCGGATGGCGGAGCTGAATTACATTCGCAACCGGTGGCAC 4571
QY 241 AACAACTGCGGGCAAAACAGTGTGCTGATTGGGGTTGCCACCTCCAGTCTGGCCCTGCG 300
DB 4570 AACAACTGCGGGCAAAACAGTGTGCTGATTGGGGTTGCCACCTCCAGTCTGGCCCTGCG 4511
QY 301 ACGGCGCGTCGCAAAATTGTTCGGCGCGATTAATAATCTCGCGCGGATCAACTGGGTGCCAGCG 360

RESULT 10
PCT-US96-05611A-21/c
; Sequence 21, Application PC/TUS9605611A
; GENERAL INFORMATION:
; APPLICANT: Mueller, John P.
; APPLICANT: Leonardo, Michael J.
; APPLICANT: McFarland, Henry F.
; APPLICANT: Matis, Louis A.
; APPLICANT: Mueller, Eileen Elliott

Db 4510 ACGCGCGTTCGCAAAATTGTTCGGCGGATTAATAATCTCGCGCGATCAACTGGGTGCACGG 4451
QY 361 TGGTGGTGTTCGATGTAAGCGGCGCTGGAAGCGCTGTAAGCGGCGGTGCAACAATC 420
DB 4450 TGGTGGTGTTCGATGTAAGCGGCGCTGGAAGCGCTGTAAGCGGCGGTGCAACAATC 4391
QY 421 TTCTCGCGCAACGGCTCAGTGGGCTGATTAATATATATATATATATATATATATATATAT 480
DB 4390 TTCTCGCGCAACGGCTCAGTGGGCTGATTAATATATATATATATATATATATATATATAT 4331
QY 481 TTGCTGTGGAAGCTGCCCTGACATTAATGTTCGGCGGTATATTTCTTGATGTCCTGACCA 540
DB 4330 TTGCTGTGGAAGCTGCCCTGACATTAATGTTCGGCGGTATATTTCTTGATGTCCTGACCA 4271
QY 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGACTGGCGCTGGAGCATC 600
DB 4270 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGACTGGCGCTGGAGCATC 4211
QY 601 TGGTGGCATTTGGGTCAACAGCAAAATCGCGCTGTGTAGCGGGCCCATTAAGTTCTGTCTCG 660
DB 4210 TGGTGGCATTTGGGTCAACAGCAAAATCGCGCTGTGTAGCGGGCCCATTAAGTTCTGTCTCG 4151
QY 661 CGCGCTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 4150 CGCGCTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 4091
QY 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTTCAACAAACCATGCAAAATGCTGA 780
DB 4090 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTTCAACAAACCATGCAAAATGCTGA 4031
QY 781 ATGAGGGCATCGTTCCCACTCGCATGTGTTGCGCAACGATCAGATGCGCTGGCGCGCAA 840
DB 4030 ATGAGGGCATCGTTCCCACTCGCATGTGTTGCGCAACGATCAGATGCGCTGGCGCGCAA 3971
QY 841 TGGCGGCAATTACCGAGTCCGGGCTGGCGTGGTGGCGATATCTCGGTAGTGGGATAG 900
DB 3970 TGGCGGCAATTACCGAGTCCGGGCTGGCGTGGTGGCGATATCTCGGTAGTGGGATAG 3911
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACCAACCATCAAAAGGATTTTC 960
DB 3910 ACGATACCGAAGACAGCTCATGTTATATCCCGCGCTTAACCAACCATCAAAAGGATTTTC 3851
QY 961 GCCTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
DB 3850 GCCTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 3791
QY 1021 AGGCAATCAGCTGTGGCGCTCTCACTGGTGAAGAAAGAAACCAACCTGGGGCCCAATA 1080
DB 3790 AGGCAATCAGCTGTGGCGCTCTCACTGGTGAAGAAAGAAAGAAACCAACCTGGGGCCCAATA 3731
QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGCAGCTGGCACGACAGGTTT 1140
DB 3730 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAATGCAGCTGGCACGACAGGTTT 3671
QY 1141 CCCGACTGGAAGCGGGCAGTGCAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
DB 3670 CCCGACTGGAAGCGGGCAGTGCAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 3611
QY 1201 GCAC 1204
DB 3610 GCAC 3607

; TITLE OF INVENTION: gp41 antigen
; FILE REFERENCE: 01-1692-A
; CURRENT APPLICATION NUMBER: US/10/263,103
; CURRENT FILING DATE: 2002-10-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 5312
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: plasmid pET-cer
US-10-263-103-35

Query Match 36.3%; Score 1199.2; DB 4; Length 5312;
Best Local Similarity 99.8%; Pred. NO. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY	1	CCGACACCATCGAATGGTGC	AAACCTTTCGGGTATGGCATGATAGCCCGGAGAGA	60
DB	4622	CGGACACCATCGAATGGCGCA	AAACCTTTCGGGTATGGCATGATAGCCCGGAGAGA	4563
QY	61	GTCAAATTCAGGGTGGTGA	ATGTGAACCAAGTAACGTTATACGATGTCGACAGTATGCCG	120
DB	4562	GTCAAATTCAGGGTGGTGA	ATGTGAACCAAGTAACGTTATACGATGTCGACAGTATGCCG	4503
QY	121	GTGTCTCTTATCAGACCG	TTTCCCGGTGGTGAACCAAGCGACGACGTTTCTGCGAAAA	180
DB	4502	GTGTCTCTTATCAGACCG	TTTCCCGGTGGTGAACCAAGCGACGACGTTTCTGCGAAAA	4443
QY	181	CGCGGAAAGTGAAGCGG	CGATGGCGAGCTGAATTACATTCCTCCACCGGTGGCAC	240
DB	4442	CGCGGAAAGTGAAGCGG	CGATGGCGAGCTGAATTACATTCCTCCACCGGTGGCAC	4383
QY	241	AACAACTCGCGGCAAA	CAGATGCTGTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300
DB	4382	AACAACTCGCGGCAAA	CAGATGCTGTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	4323
QY	301	ACGCGCCGTCGCAAA	TTTCGCGCGGATTAATCTCGCGCGATCAACTGGGTGCACGC	360
DB	4322	ACGCGCCGTCGCAAA	TTTCGCGCGGATTAATCTCGCGCGATCAACTGGGTGCACGC	4263
QY	361	TGGTGGTGTGATGTA	GAACGAAGCGCGTGAAGCCTGTAAAGCGGCGGTGCAATC	420
DB	4262	TGGTGGTGTGATGTA	GAACGAAGCGCGTGAAGCCTGTAAAGCGGCGGTGCAATC	4203
QY	421	TTCTCGCGCAACGCT	CAGTGGCTGATCATTAATCTCCGTGGATGACCAAGATGCCA	480
DB	4202	TTCTCGCGCAACGCT	CAGTGGCTGATCATTAATCTCCGTGGATGACCAAGATGCCA	4143
QY	481	TTGCTGTGGAAGCT	GCCTGCACATAATGTTCCGGCGTTATTTCTGATGTCTCTGAC	540
DB	4142	TTGCTGTGGAAGCT	GCCTGCACATAATGTTCCGGCGTTATTTCTGATGTCTCTGAC	4083
QY	541	CACCCATCAACAGTA	TATTTCTCCATGAAGCGGTACGGCATGGGCGTGGAGCATC	600
DB	4082	CACCCATCAACAGTA	TATTTCTCCATGAAGCGGTACGGCATGGGCGTGGAGCATC	4023
QY	601	TGGTGCATTTGGTC	CACAGCAAAATCGCGCTGTAGCGGCGCCATTAAAGTTCTGTC	660
DB	4022	TGGTGCATTTGGTC	CACAGCAAAATCGCGCTGTAGCGGCGCCATTAAAGTTCTGTC	3963
QY	661	CGCGTCTCGCTCG	CTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGAT	720
DB	3962	CGCGTCTCGCTCG	CTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGAT	3903
QY	721	CGGAAACGGGAAG	CGACTGGAGTGCATGTCGGTTCCTCAACAAACCATGCAAAATGCT	780
DB	3902	CGGAAACGGGAAG	CGACTGGAGTGCATGTCGGTTCCTCAACAAACCATGCAAAATGCT	3843
QY	781	ATGAGGGCATCGT	TCCCACTCGATGCTGGTTGCCAAGCATGAGTGGCGTGGCGGCA	840
DB	3842	ATGAGGGCATCGT	TCCCACTCGATGCTGGTTGCCAAGCATGAGTGGCGTGGCGGCA	3783

QY	841	TGCGCGCATTTAC	CGAGTCCGGGCTCGCGTTGGTGGGATATCTCGTAGTGGATACG	900
DB	3782	TGCGCGCATTTAC	CGAGTCCGGGCTCGCGTTGGTGGGATATCTCGTAGTGGATACG	3723
QY	901	ACGATACCGAAG	ACAGCTCATGTTTATATCCCGCGTTAAACCAACCATCAAAACGAGTTTC	960
DB	3722	ACGATACCGAAG	ACAGCTCATGTTTATATCCCGCGTTAAACCAACCATCAAAACGAGTTTC	3663
QY	961	GCCTGCTGGG	CAAAACAGCGTGCACCTTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	1020
DB	3662	GCCTGCTGGG	CAAAACAGCGTGCACCTTTGCTGCAACTCTCTCAGGGCCAGCGGTGA	3603
QY	1021	AGGCAATCAG	CTGTGGCCGCTCTCACTGGTGAAGAAACCAACCTTGGCGCCCAATA	1080
DB	3602	AGGCAATCAG	CTGTGGCCGCTCTCACTGGTGAAGAAACCAACCTTGGCGCCCAATA	3543
QY	1081	CGAAACCGC	CTCTCCCGCGTTGGCGATTCAATTAATCAGCTGGCAGCAGAGTTT	1140
DB	3542	CGAAACCGC	CTCTCCCGCGTTGGCGATTCAATTAATCAGCTGGCAGCAGAGTTT	3483
QY	1141	CCGACTGGAA	AGCGGCGAGTGAGCGCAACCAATTAATGTAGTTAGCTCACTCATTTAG	1200
DB	3482	CCGACTGGAA	AGCGGCGAGTGAGCGCAACCAATTAATGTAGTTAGCTCACTCATTTAG	3423
QY	1201	GCAC	1204	
DB	3422	GCAC	3419	

RESULT 12
US-08-929-967-1/c
; Sequence 1, Application US/08929967
; Patent No. 5891637
; GENERAL INFORMATION:
; APPLICANT: Ruppert, Siegfried J.W.
; TITLE OF INVENTION: Construction of Full-length cDNA Libraries
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/929,967
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: P1035R1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-3216
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5443 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
US-08-929-967-1

Query Match 36.3%; Score 1199.2; DB 2; Length 5443;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAGAACTTTCCGGGTATGGCATATAGCGCCCGGAAGAGA 60
Db 4810 CGGACACCATCGAATGGGCGCAAACTTTCCGGGTATGGCATATAGCGCCCGGAAGAGA 4751
QY 61 GTCATTCAGGGTGGTGAATGTGAACCGATGAACGTTATACGATTCGCGAGTATGCCG 120
Db 4750 GTCATTCAGGGTGGTGAATGTGAACCGATGAACGTTATACGATTCGCGAGTATGCCG 4691
QY 121 GTGTCTCTTATCAGACCGGTTTCCCGCGTGGTGAACCGCGCAGCCACGTTTCTGCGAAA 180
Db 4690 GTGTCTCTTATCAGACCGGTTTCCCGCGTGGTGAACCGCGCAGCCACGTTTCTGCGAAA 4631
QY 181 CCGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 240
Db 4630 CCGCGGAAAAAGTGAAGCGCGATGGCGGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 4571
QY 241 AACAACTCGCGGCAAAAGTCGTGCTGATTTGGGTTGCCACCTCAGTCTGGCCCTGC 300
Db 4570 AACAACTCGCGGCAAAAGTCGTGCTGATTTGGGTTGCCACCTCAGTCTGGCCCTGC 4511
QY 301 ACGCGCCGTGCGAAATTTGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCGACG 360
Db 4510 ACGCGCCGTGCGAAATTTGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCGACG 4451
QY 361 TGGTGGTGCATGTGAAGCAAGCGCGCTGGAAGCTGTAAAGCGCGGTGCACAATC 420
Db 4450 TGGTGGTGCATGTGAAGCAAGCGCGCTGGAAGCTGTAAAGCGCGGTGCACAATC 4391
QY 421 TTCTCGCGCAACGGTCACTGGGCTGATTAATTAATCTCGCGCGATCAACTGGGTGCGAC 480
Db 4390 TTCTCGCGCAACGGTCACTGGGCTGATTAATTAATCTCGCGCGATCAACTGGGTGCGAC 4331
QY 481 TTGTGTGGAAGTGCCTGCACTAATGTTCCGGGTTATTTCTTGATGCTCTGACAGAGA 540
Db 4330 TTGTGTGGAAGTGCCTGCACTAATGTTCCGGGTTATTTCTTGATGCTCTGACAGAGA 4271
QY 541 CACCCATCAACAGTATTAATTTTCCCATGAAGACGGTACCGCATCGGCGGTGAGCATC 600
Db 4270 CACCCATCAACAGTATTAATTTTCCCATGAAGACGGTACCGCATCGGCGGTGAGCATC 4211
QY 601 TGGTGGCATTTGGGTCAACAGCAATCGCGCTGTTAGCGGGCCCATTAAGTCTCTCTCGG 660
Db 4210 TGGTGGCATTTGGGTCAACAGCAATCGCGCTGTTAGCGGGCCCATTAAGTCTCTCTCGG 4151
QY 661 CGGCTCTGCGTCTGGCTGGGCAATAATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db 4150 CGGCTCTGCGTCTGGCTGGGCAATAATCTCACTCGCAATCAAAATTCAGCGGATAG 4091
QY 721 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTCAACAAACCATGCAATGCTGA 780
Db 4090 CGGAACGGGAAGGCGACTGGAGTGCATGTCGGGTTTTCACAAACCATGCAATGCTGA 4031
QY 781 ATGAGGGCATCGTTCCCACTGCGATGCTGGTTCGCAACGATCAGATGGCGCTGGGCGCAA 840
Db 4030 ATGAGGGCATCGTTCCCACTGCGATGCTGGTTCGCAACGATCAGATGGCGCTGGGCGCAA 3971
QY 841 TGCGCGCATTCAGAGTCCGGGCTCGCGTGGTGGTGGATATCTCGGTAGTGGGATACG 900
Db 3970 TGCGCGCATTCAGAGTCCGGGCTCGCGTGGTGGTGGATATCTCGGTAGTGGGATACG 3911
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCATCAACACGATTTTC 960
Db 3910 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTTAAACCATCAACACGATTTTC 3851
QY 961 GCCTGTGGGGCAAAACAGCGTGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 3850 GCCTGTGGGGCAAAACAGCGTGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 3791
QY 1021 AGGCAATCAGCTGTTCCCGCTCTCACTGTTGAAAGAAAAACCACTCGCGGCCCAATA 1080
Db 3790 AGGCAATCAGCTGTTCCCGCTCTCACTGTTGAAAGAAAAACCACTCGCGGCCCAATA 3731
QY 1081 CGCAACCGGCTCTCTCCCGCGCGTGGCGGATTCATTAATGCAGTGGCACGACAGTTT 1140

Db 3730 CGCAACCGGCTCTCTCCCGCGGTTGGCGGATTCATTAATGCAGTGGCACGACAGTTT 3671
QY 1141 CCGACTGGAAAGCGGCGAGTGAAGCGCAACGCAATTAATGTAGCTTACCTCATTAG 1200
Db 3670 CCGACTGGAAAGCGGCGAGTGAAGCGCAACGCAATTAATGTAGCTTACCTCATTAG 3611
QY 1201 GCAC 1204
Db 3610 GCAC 3607
RESULT 13
US-09-702-705-785/c
; Sequence 785, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCES: 210121.478C14
; CURRENT APPLICATION NUMBER: US/09/702,705
; CURRENT FILING DATE: 2000-10-30
; NUMBER OF SEQ ID NOS: 1833
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-702-705-785
Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CCGACACCATCGAATGGTGCAGAACTTTCCGGGTATGGCATATAGCGCCCGGAAGAGA 60
Db 4677 CGGACACCATCGAATGGGCGCAAACTTTCCGGGTATGGCATATAGCGCCCGGAAGAGA 4618
QY 61 GTCATTCAGGGTGGTGAATGTGAACCGATGAACGTTATACGATTCGCGAGATATGCCG 120
Db 4617 GTCATTCAGGGTGGTGAATGTGAACCGATGAACGTTATACGATTCGCGAGATATGCCG 4558
QY 121 GTGTCTCTTATCAGACCGGTTTCCCGCGTGGTGAACCGCGCAGCCACGTTTCTGCGAAA 180
Db 4557 GTGTCTCTTATCAGACCGGTTTCCCGCGTGGTGAACCGCGCAGCCACGTTTCTGCGAAA 4498
QY 181 CCGCGGAAAAAGTGAAGCGGCGATGGCGGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 240
Db 4497 CCGCGGAAAAAGTGAAGCGGCGATGGCGGAGCTGAATTTACATTCCTCAACCGCGTGGCAC 4438
QY 241 AACAACTCGCGGCAAAAGTCGTGCTGATTTGGGTTGCCACCTCAGTCTGGCCCTGC 300
Db 4437 AACAACTCGCGGCAAAAGTCGTGCTGATTTGGGTTGCCACCTCAGTCTGGCCCTGC 4378
QY 301 ACGCGCGCTCGAAATTTGTCGCGCGGATTAATCTCGCGCGATCAACTGGGTGCGCAG 360
Db 4377 ACGCGCGCTCGAAATTTGTCGCGCGGATTAATCTCGCGCGATCAACTGGGTGCGCAG 4318
QY 361 TGGTGGTGCATGTGAAGCAAGCGCGTGAAGACCTGTAAAGCGCGGTGCACAATC 420
Db 4317 TGGTGGTGCATGTGAAGCAAGCGCGTGAAGACCTGTAAAGCGCGGTGCACAATC 4258
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGCTGGATGACCGAGATGCCA 480

Db 4257 TTCTCGCAACGCGTCAGTGGCTGATCATTAATCTATCCGTCGATGATGACGAGTGC 4198
Qy 481 TTGCTGTGGAAGTCCCTGACATAAGTTTCGGCGCTTATTTCTTGATGTCTGACGAGA 540
Db 4197 TTGCTGTGGAAGTCCCTGACATAAGTTTCGGCGCTTATTTCTTGATGTCTGACGAGA 4138
Qy 541 CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 600
Db 4137 CACCCATCAACAGTATTATTTCTCCCATGAAGCGGTACGCGACTGGCGGTGGAGCATC 4078
Qy 601 TGGTGCATTTGGGTCCACGACAAATCGCGCTGTTAGCGGGGCCCAATAAGTTCTGTCTCGG 660
Db 4077 TGGTGCATTTGGGTCCACGACAAATCGCGCTGTTAGCGGGGCCCAATAAGTTCTGTCTCGG 4018
Qy 661 CGCGTCTGGCTGTGGCTGGCGATAAATATCTCACTCGGATCAAAATTCAGCCGATAG 720
Db 4017 CGCGTCTGGCTGTGGCTGGCGATAAATATCTCACTCGGATCAAAATTCAGCCGATAG 3958
Qy 721 CGGAACGGGAAGCGGACCTGGAGTGCATGTCGGGTTTCAACAAACCATGCAATGCTGA 780
Db 3957 CGGAACGGGAAGCGGACCTGGAGTGCATGTCGGGTTTCAACAAACCATGCAATGCTGA 3898
Qy 781 ATGAGGGCATCGTTCCTCCACTGCGATGCTGGTTGCCAAGCATCAGATGGCGCTGGCGCAA 840
Db 3897 ATGAGGGCATCGTTCCTCCACTGCGATGCTGGTTGCCAAGCATCAGATGGCGCTGGCGCAA 3838
Qy 841 TGC CGCGCATTAACGAGTCCGGGCTGGCGGTTGGTGGCGGATATCTCGGTAGTGGGATACG 900
Db 3837 TGC CGCGCATTAACGAGTCCGGGCTGGCGGTTGGTGGCGGATATCTCGGTAGTGGGATACG 3778
Qy 901 ACGATACCGAAGACAGCTCATGTTATNTCCCGCGGTTAACCACCATCAACAGGATTTTC 960
Db 3777 ACGATACCGAAGACAGCTCATGTTATNTCCCGCGGTTAACCACCATCAACAGGATTTTC 3718
Qy 961 GCCTGCTGGGCAACACGAGCTGGACCGCTTCTGCAACTCTCTCAGGCGCAGGCGGTGA 1020
Db 3717 GCCTGCTGGGCAACACGAGCTGGACCGCTTCTGCAACTCTCTCAGGCGCAGGCGGTGA 3658
Qy 1021 AGGCAATCAGCTGTGGCGGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGGCCAATA 1080
Db 3657 AGGCAATCAGCTGTGGCGGCTCTCACTGGTGAAGAAAGAAACCAACCTGGCGGCCAATA 3598
Qy 1081 CGCAACCGCTCTCCCGCGGCTTGGCGGATTCATTATGCGAGCTGGCAGCAGAGTTT 1140
Db 3597 CGCAACCGCTCTCTCCCGCGGCTTGGCGGATTCATTATGCGAGCTGGCAGCAGAGTTT 3538
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 1200
Db 3537 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGTCACTCATTTAG 3478
Qy 1201 GCAC 1204
Db 3477 GCAC 3474

RESULT 14

US-09-736-457-785/c
; Sequence 785, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Panger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15

; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-785

Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGGCAAAACCTTTTCGGGTATGGCATGATAGCCCGGAAGAGA 60
Db 4677 CCGACACCATCGAATGGCGCAAAACCTTTTCGGGTATGGCATGATAGCCCGGAAGAGA 4618
Qy 61 GTCATATCAGGCTGGTGAATCTGAAACCCAGTAACGTTATACGATCTCGCAGATATGCCG 120
Db 4617 GTCATATCAGGCTGGTGAATCTGAAACCCAGTAACGTTATACGATCTCGCAGATATGCCG 4558
Qy 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCCAGCCAGCCACCGTTTCTGCAAAA 180
Db 4557 GTGTCTCTTATCAGACCGTTTCCCGCGTGGTGAACCCAGCCAGCCACCGTTTCTGCAAAA 4498
Qy 181 CGCGGAAAAAGTGGAGCGCGCATGGCGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
Db 4497 CGCGGAAAAAGTGGAGCGCGCATGGCGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 4438
Qy 241 AACAACTGGCGGGCAACAGTCTGTGATTTGGGTTGGCCACCTCCAGTCTGGCCCTGC 300
Db 4437 AACAACTGGCGGGCAACAGTCTGTGATTTGGGTTGGCCACCTCCAGTCTGGCCCTGC 4378
Qy 301 ACGCGCGTCTGCAAAATCTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGGCAGCG 360
Db 4377 ACGCGCGTCTGCAAAATCTCGCGCGGATTAATCTCGCGCGGATCAACTGGGTGGCAGCG 4318
Qy 361 TGGTGGTTCGATGGTGAAGCAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCAACAATC 420
Db 4317 TGGTGGTTCGATGGTGAAGCAAGCGCGCTCGAAGCCTGTAAAGCGCGGTGCAACAATC 4258
Qy 421 TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATCTCCGTGGATCAGCAGGATGCA 480
Db 4257 TTCTCGCGCAACGCGTCAGTGGGCTGATCATTAATCTCCGTGGATCAGCAGGATGCA 4198
Qy 481 TTGTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTATTTCTTGATGTCTCTGACCCAGA 540
Db 4197 TTGTGTGGAAGCTGCCTGCACTAATGTTCCGGCGTATTTCTTGATGTCTCTGACCCAGA 4138
Qy 541 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGGTGGAGCATC 600
Db 4137 CACCCATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGGTGGAGCATC 4078
Qy 601 TGGTGCATTTGGGTCCACGACAAATCGCGCTGTTAGCGGGGCCCAATAAGTTCTGTCTCGG 660
Db 4077 TGGTGCATTTGGGTCCACGACAAATCGCGCTGTTAGCGGGGCCCAATAAGTTCTGTCTCGG 4018
Qy 661 CGCGTCTGGCTGTGGCTGGCGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 4017 CGCGTCTGGCTGTGGCTGGCGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 3958
Qy 721 CGGAACGGGAAGCGGACCTGGAGTGCATGTCGGGTTTCAACAAACCATGCAATGCTGA 780
Db 3957 CGGAACGGGAAGCGGACCTGGAGTGCATGTCGGGTTTCAACAAACCATGCAATGCTGA 3898
Qy 781 ATGAGGGCATCGTTCCTCCACTGCGATGCTGGTTGCCAAGCATCAGATGGCGCTGGCGCAA 840
Db 3897 ATGAGGGCATCGTTCCTCCACTGCGATGCTGGTTGCCAAGCATCAGATGGCGCTGGCGCAA 3838
Qy 841 TGC CGCGCATTAACGAGTCCGGGCTGGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 3837 TGC CGCGCATTAACGAGTCCGGGCTGGCGGTTGGTGGGATATCTCGGTAGTGGGATACG 3778


```
QY 901 ACATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACACCATCAAAACAGGATTTTC 960
Db 3777 ACATACCGAAGACAGCTCATGTTATATCCCGCCGTTAAACACCATCAAAACAGGATTTTC 3718
QY 961 GCCTGCTGGGGCAAAACAGCGTGGACCGGTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 3717 GCCTGCTGGGGCAAAACAGCGTGGACCGGTTGCTGCAACTCTCTCAGGGCCAGGCGGTGA 3658
QY 1021 AGGGCAATCAGCTGTTGCCCGTCTCAGTGGTGAAAGAAACCAACCCCTGGGGCCCAATA 1080
Db 3657 AGGGCAATCAGCTGTTGCCCGTCTCAGTGGTGAAAGAAACCAACCCCTGGGGCCCAATA 3598
QY 1081 CGAAACCCCTCTCCCGCGGTTGGCGGATTCATTAAATGAGCTGGCAGCAGCAGGTTT 1140
Db 3597 CGAAACCCCTCTCCCGCGGTTGGCGGATTCATTAAATGAGCTGGCAGCAGCAGGTTT 3538
QY 1141 CCCGACTGGAAAGCGGCGAGTGAGGCGAAACCGAATTAATGAGTTAGCTCACTCATTTAG 1200
Db 3537 CCCGACTGGAAAGCGGCGAGTGAGGCGAAACCGAATTAATGAGTTAGCTCACTCATTTAG 3478
QY 1201 GCAC 1204
Db 3477 GCAC 3474

RESULT 15
US-09-614-124B-785/c
; Sequence 785, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedrick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614.124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 785
; LENGTH: 5502
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-785

Query Match 36.3%; Score 1199.2; DB 4; Length 5502;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1201; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGTTGCAAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 4677 CCGACACCATCGAATGTTGCAAAACCTTTCCGGTATGGCATGATAGCGCCCGGAAGAGA 4618
QY 61 GTCAATTCAGGGTGGTGAATGTGAACAGTAAAGCTTATAGATGTCGCGAGATGCGG 120
Db 4617 GTCAATTCAGGGTGGTGAATGTGAACAGTAAAGCTTATAGATGTCGCGAGATGCGG 4558
QY 121 GTGTCTCTTATCAGACCGGTTTCCCGCGTGGTGAACACGAGCCAGCAGCTTCTGCGAAA 180
Db 4557 GTGTCTCTTATCAGACCGGTTTCCCGCGTGGTGAACACGAGCCAGCAGCTTCTGCGAAA 4498
QY 181 CGCGGAAAAAGTGAAGCGGCGATGCGGAGCTGGAATTAATTCACACCGCGTGGCAC 240
Db 4497 CGCGGAAAAAGTGAAGCGGCGATGCGGAGCTGGAATTAATTCACACCGCGTGGCAC 4438
QY 241 AACCACTGGCGGGCAACAGTCGTTGCTGATTGGCGTTGCCACTCCAGTCTGGCCCTGC 300
```

Search completed: August 3, 2005, 02:32:31.
Job time : 356.4 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 1309.2 Seconds

(without alignments)
16314.373 Million cell updates/sec

Title: US-09-765-555B-17
Perfect score: 3300
Sequence: 1 ccgcacacatcgatgtgc.....acgacgttcaggactagcct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3300	100.0	3300	10	US-09-765-555-17
2	3232.8	98.0	3300	10	US-09-765-555-18
3	3215.2	97.4	3300	10	US-09-765-555-14
4	3215.2	97.4	3300	10	US-09-765-555-15
5	3207.2	97.2	3300	10	US-09-765-555-16
6	2705.8	82.0	7373	17	US-10-257-384A-1
7	2705.8	82.0	8101	17	US-10-149-472-5

8	2701	81.8	6648	19	US-10-343-859-8	Sequence 8, Appli
9	2698	81.8	9191	19	US-10-343-859-9	Sequence 9, Appli
10	2690.4	81.5	7112	18	US-10-263-153-40	Sequence 40, Appli
11	2690.4	81.5	7259	18	US-10-263-153-30	Sequence 30, Appli
12	2690.4	81.5	7322	18	US-10-263-153-35	Sequence 35, Appli
13	2690.4	81.5	7352	18	US-10-263-153-25	Sequence 25, Appli
14	2690.4	81.5	7370	18	US-10-263-153-20	Sequence 20, Appli
15	2690.4	81.5	7370	18	US-10-263-153-61	Sequence 61, Appli
16	2690.4	81.5	7370	18	US-10-263-153-66	Sequence 66, Appli
17	2690.4	81.5	7370	18	US-10-263-153-71	Sequence 71, Appli
18	2690.4	81.5	7403	18	US-10-263-153-15	Sequence 15, Appli
19	2690.4	81.5	7442	18	US-10-263-153-10	Sequence 10, Appli
20	2690.4	81.5	7478	18	US-10-263-153-3	Sequence 3, Appli
21	2636.8	79.9	6806	10	US-09-908-483A-194	Sequence 194, App
22	2636.8	79.9	6806	20	US-10-801-487-194	Sequence 194, App
23	2636.8	79.9	6806	20	US-10-801-938-194	Sequence 194, App
24	2636.8	79.9	6806	20	US-10-801-509-194	Sequence 194, App
25	2636.8	79.9	6806	20	US-10-801-486-194	Sequence 194, App
26	2636.8	79.9	6806	21	US-10-801-493-194	Sequence 194, App
27	2605.4	79.0	7553	18	US-10-263-153-7	Sequence 7, Appli
28	1609.4	48.8	5558	16	US-10-241-596-137	Sequence 137, App
29	1474	44.7	4700	17	US-10-313-963A-54	Sequence 54, Appli
30	1462.4	44.3	4921	21	US-10-495-491-2	Sequence 2, Appli
31	1462.4	44.3	4935	21	US-10-495-491-1	Sequence 1, Appli
32	1462.4	44.3	4945	21	US-10-495-491-3	Sequence 3, Appli
33	1462.4	44.3	4951	21	US-10-495-491-4	Sequence 4, Appli
34	1417.6	43.0	1511	10	US-09-987-763-30	Sequence 30, Appli
35	1417.6	43.0	5926	13	US-10-024-809-3	Sequence 3, Appli
36	1238.4	37.5	5024	17	US-10-313-963A-52	Sequence 52, Appli
37	1238.4	37.5	5448	21	US-10-671-995A-2	Sequence 2, Appli
38	1238.4	37.5	5454	21	US-10-671-995A-1	Sequence 1, Appli
39	1237.2	37.5	1922	10	US-09-995-988A-32	Sequence 32, Appli
40	1237.2	37.5	1922	10	US-10-420-034A-32	Sequence 32, Appli
41	1237.2	37.5	6270	21	US-10-855-897-2	Sequence 2, Appli
42	1204	36.5	4557	17	US-10-457-372-5	GENERAL INFORMA
43	1204	36.5	6664	18	US-10-270-176-1	Sequence 1, Appli
44	1204	36.5	6668	18	US-10-270-176-2	Sequence 2, Appli
45	1204	36.5	6742	18	US-10-270-176-19	Sequence 19, Appli

ALIGNMENTS

RESULT 1
US-09-765-555-17
; Sequence 17, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m4 and zinc finger
; OTHER INFORMATION: protein 2Fpm4
US-09-765-555-17

Query Match 100.0%; Score 3300; DB 10; Length 3300;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3300; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCGGAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCGGAGAGA 60
QY 61 GTCAATTCAGGTTGGTGAATGTGAACACCAAGTAACTGATACGATGTGCGAGATATGCCG 120
Db 61 GTCAATTCAGGTTGGTGAATGTGAACACCAAGTAACTGATACGATGTGCGAGATATGCCG 120
QY 121 GTGTCTTTATCAGACCGTTTCCGGGTGGTGAACCAAGCCAGCCAGCGTTTCTGGGAAA 180
Db 121 GTGTCTTTATCAGACCGTTTCCGGGTGGTGAACCAAGCCAGCCAGCGTTTCTGGGAAA 180
QY 181 CCGGGGAAAAAGTGAAGCGCGGATGCGGAGCTGAATTTACATTTCCCAACCGCTGGCAC 240
Db 181 CCGGGGAAAAAGTGAAGCGCGGATGCGGAGCTGAATTTACATTTCCCAACCGCTGGCAC 240
QY 241 AACAACTGGCGGCAAAACAGTCTGCTGATTTGGGTTGGCCACTTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGCAAAACAGTCTGCTGATTTGGGTTGGCCACTTCCAGTCTGGCCCTGC 300
QY 301 AGCGCCGTCGAAATTTGTCGGCGGATTAATCTCGGCCGATCAACTGGGTGCCAGCG 360
Db 301 AGCGCCGTCGAAATTTGTCGGCGGATTAATCTCGGCCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGTGTCGATGCTGAGAACGAAGCGGCTCGAAGCCTGTAAGCGGCGGTGCAATC 420
Db 361 TGGTGTGTCGATGCTGAGAACGAAGCGGCTCGAAGCCTGTAAGCGGCGGTGCAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTCGCTGGATCACCAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCATTAATCTCGCTGGATCACCAGGATGCCA 480
QY 481 TTGCTGTGAAGCTGCTGCACTAATGTCGGCGTTATTTCTTGATGCTCTGACCAGA 540
Db 481 TTGCTGTGAAGCTGCTGCACTAATGTCGGCGTTATTTCTTGATGCTCTGACCAGA 540
QY 541 CACCCATCAACAGTATTTTCTCCCATGAGAGCGTACCGGCTGGGCTGGGAGATC 600
Db 541 CACCCATCAACAGTATTTTCTCCCATGAGAGCGTACCGGCTGGGCGTGGAGATC 600
QY 601 TGGTGCATTTGGTCCACGAAATCGGCTGTTAGCGGGGCCATTAAGTTCTGCTCGG 660
Db 601 TGGTGCATTTGGTCCACGAAATCGGCTGTTAGCGGGGCCATTAAGTTCTGCTCGG 660
QY 661 CGGCTCGGCTGCTGGCTGGCATTAATCTCACTCGCAATCAAATTCAGCCGATAG 720
Db 661 CGGCTCGGCTGCTGGCTGGCATTAATCTCACTCGCAATCAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGCGGACCTGGAGTGCATGTCGGTTCGCAACGATCAGATGCGCTGGCGCAA 840
Db 721 CGGAACGGGAAGCGGACCTGGAGTGCATGTCGGTTCGCAACGATCAGATGCGCTGGCGCAA 840
QY 841 TGGCGCCATTTACAGTTCGGGCTGGGCTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGCCATTTACAGTTCGGGCTGGGCTGGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAACAGGATTTTC 960
Db 901 ACGATACGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACGAGTGAACGCTTGTGCAACTCTCGAGGCGACGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGTGAACGCTTGTGCAACTCTCTCAGGCGCAGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTCGCGCTCTCACTGTGAAAGAAAACACCCCTGGCGCCCAATA 1080
Db 1021 AGGCAATCAGCTGTCGCGCTCTCACTGTGAAAGAAAACACCCCTGGCGCCCAATA 1080
QY 1081 CGMAACCGGCTCTCCCGCGGCTTGGCGGATTCATTTAATGCGAGTGGCACGAGGTTT 1140

Db 1081 CGMAACCGGCTCTCCCGCGGCTTGGCGGATTCATTTAATGCGAGTGGCACGAGGTTT 1140
QY 1141 CCGCATGGAAGCGGCGAGTGAGCGAAACCAATTAATGTAGTGTAGCTCACCTCATTTAG 1200
Db 1141 CCGGACTGGAAGCGGCGAGTGAGCGAAACCAATTAATGTAGTGTAGCTCACCTCATTTAG 1200
QY 1201 GCACAAATCTCATGTTTGAACAGCTTATCATGCACTGCAAGCGTCAACAAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGAACAGCTTATCATGCACTGCAAGCGTCAACAAATGCTTCTGGCG 1260
QY 1261 TCAGGACGCAATCGGAAAGCTGTGATGCGTGTGCGAGCTGTAATCACTGCAATTAATTCG 1320
Db 1261 TCAGGACGCAATCGGAAAGCTGTGATGCGTGTGCGAGCTGTAATCACTGCAATTAATTCG 1320
QY 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTTTTGGCGCGACATCATAAACGGTT 1380
Db 1321 TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTTTTGGCGCGACATCATAAACGGTT 1380
QY 1381 CTGGCAATATCTGAATGAGCTGTGACAAATTAATCATCTGGCTCGTATAATGTGTGGA 1440
Db 1381 CTGGCAATATCTGAATGAGCTGTGACAAATTAATCATCTGGCTCGTATAATGTGTGGA 1440
QY 1441 ATTGTGAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAAAGTGTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAAAGTGTTCACGA 1500
QY 1501 GCATTTACCAACAGGACCATAGATATGAAACTGAAGAGTAACTGGTAAATCTGG 1560
Db 1501 GCATTTACCAACAGGACCATAGATATGAAACTGAAGAGTAACTGGTAAATCTGG 1560
QY 1561 ATTAAACCGGATTAAGGCTATAACGGTCTCGCTGAAAGTGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAAACCGGATTAAGGCTATAACGGTCTCGCTGAAAGTGGTAAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGTAAACTGGAAGAGAAATTCGCCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGTAAACTGGAAGAGAAATTCGCCACAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATTTCTTCTGGGCAACAGCCGTTTGGTGCTAC 1740
Db 1681 GCGGCAACTGGCGATGGCCCTGACATTTCTTCTGGGCAACAGCCGTTTGGTGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGCTGAAATCAACCGGCAAGAGCTTCCAGACAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGCTGAAATCAACCGGCAAGAGCTTCCAGACAGCTGTAT 1800
QY 1801 CGTTCCTGGATGCGCTGATTAACCGCAAGCTGATTTACCGGATCGCTGTT 1860
Db 1801 CGTTCCTGGATGCGCTGATTAACCGCAAGCTGATTTACCGGATCGCTGTT 1860
QY 1861 GAAGCGTATCGCTGATTTTAAACAAAGATCTGCTGCCGAACCGCGCAAAACCTGGGAA 1920
Db 1861 GAAGCGTATCGCTGATTTTAAACAAAGATCTGCTGCCGAACCGCGCAAAACCTGGGAA 1920
QY 1921 GAGATCCCGGCTGGATAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCTGGATAAGAACTGAAAGCGAAAGGTAAAGAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCTGCTTACCTGGCGCTGATGCTGTGACGCGGGTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACCTGCTTACCTGGCGCTGATGCTGTGACGCGGGTTATCGGTTCAAG 2040
QY 2041 TATGAAAACGCAAGTACGACATTAAGAGCTGGGCTGGATAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAACGCAAGTACGACATTAAGAGCTGGGCTGGATAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGAACCTTCTGCTGTGAACCTGATTAATAAACAACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGAACCTTCTGCTGTGAACCTGATTAATAAACAACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAAGAGCTGCCTTTAATAAGGCGAAACAGCGATGACCATCAACGGCCCGTGG 2220

Db 2161 TCCATCGCAGAGCTGCCCTTTAATAAAGCGCAAAACAGCGATGACCATCAA CGGCCCGTGG 2220
Qy 2221 GCATGGTCCAAATCGACACACGAAAGTGAATTAATGTGTAAACGGTACTGCGCACTTC 2280
Db 2221 GCATGGTCCAAATCGACACACGAAAGTGAATTAATGTGTAAACGGTACTGCGCACTTC 2280
Qy 2281 AAGGGTCAACCATCCAAACCGTTCTGTCGGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGGTCAACCATCCAAACCGTTCTGTCGGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Qy 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
Db 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTCTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGAAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGAAAG 2460
Qy 2461 TTGGCGAAGATCCACGTTATTCGCCGCCACCATGGAAACGCCACAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCACGTTATTCGCCGCCACCATGGAAACGCCACAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGGCTTTCTGCTATGCCGTGCGTACTGCGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGGCTTTCTGCTATGCCGTGCGTACTGCGTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACAGTCTCGATGAAGCCCTCGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTCTCGATGAAGCCCTCGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Qy 2641 AACACACACAAATAAACAAACAAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Db 2641 AACACACACAAATAAACAAACAAACCTCGGGATCGAGGGAAGGATTCAGAAATTC 2700
Qy 2701 GGATCCTCTCTCTCTGTCGCGCGCCCTCGAGCCCGGGGAGAAGCCCTATCTGTT 2760
Db 2701 GGATCCTCTCTCTCTGTCGCGCGCCCTCGAGCCCGGGGAGAAGCCCTATCTGTT 2760
Qy 2761 CCGGAATGTGTAAGTCTCTTACGACAGAGCAGTCCCTGGTGCACACAGCGTACCCAC 2820
Db 2761 CCGGAATGTGTAAGTCTCTTACGACAGAGCAGTCCCTGGTGCACACAGCGTACCCAC 2820
Qy 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAATCTTTTAGCCAGAGCAGCAGC 2880
Db 2821 ACGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAATCTTTTAGCCAGAGCAGCAGC 2880
Qy 2881 CTGGTGCACATCAACGCACTCATATGCGGAGAAGCCATACAAATGTCCAGATGTGCG 2940
Db 2881 CTGGTGCACATCAACGCACTCATATGCGGAGAAGCCATACAAATGTCCAGATGTGCG 2940
Qy 2941 AAGTCTTTCAGTGATGTGCTGATCTTTCGAGGCAACCAACGTAATGTCAGAAATGTGCG 3000
Db 2941 AAGTCTTTCAGTGATGTGCTGATCTTTCGAGGCAACCAACGTAATGTCAGAAATGTGCG 3000
Qy 3001 CCCTATGCTTGTCCGGAATGTGGTAAGTCTTCTCTCAGAGCTCTCACCCTGGTGCACAC 3060
Db 3001 CCCTATGCTTGTCCGGAATGTGGTAAGTCTTCTCTCAGAGCTCTCACCCTGGTGCACAC 3060
Qy 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAATCTTTTAGC 3120
Qy 3121 CCGAGCGATAACCTGGTGGCCATCAACGCACTCATATCTGCGGAGAAGCCATACAAATGT 3180
Db 3121 CCGAGCGATAACCTGGTGGCCATCAACGCACTCATATCTGCGGAGAAGCCATACAAATGT 3180
Qy 3181 CCAGAAATGTGCAAGTCTTTTCTCAAATTCAGGCCAATTTGGTCCGTCACCAAGTACTCAC 3240
Db 3181 CCAGAAATGTGCAAGTCTTTTCTCAAATTCAGGCCAATTTGGTCCGTCACCAAGTACTCAC 3240
Qy 3241 ACCGGTAAAAAATAGTGGCCAGCGCCGCTAGTACCGGTACGAGTTCCGAGCTACCGT 3300
Db 3241 ACCGGTAAAAAATAGTGGCCAGCGCCGCTAGTACCGGTACGAGTTCCGAGCTACCGT 3300

RESULT 2
US-09-765-555-18
; Sequence 18, Application US/09765555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; FILE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Parial sequence of pMal-Ap3 and zinc finger
; OTHER INFORMATION: protein ZFPAP3
US-09-765-555-18

Query Match 98.0%; Score 3232.8; DB 10; Length 3300;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 3258; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGC AAAACCTTTCCCGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGC AAAACCTTTCCCGGTATGGCATGATAGCGCCCGGAAGAGA 60
Qy 61 GTCAATTCAGGGTGGTGAATGTGAACCCAGTAAAGTTATACATGATGTCAGAGTATGCG 120
Db 61 GTCAATTCAGGGTGGTGAATGTGAACCCAGTAAAGTTATACATGATGTCAGAGTATGCG 120
Qy 121 GTGTCTCTTATCAGACCTGTTCCCGCGTGGTGAACCCAGCCAGCCAGCTTTCTCGGAAAA 180
Db 121 GTGTCTCTTATCAGACCTGTTCCCGCGTGGTGAACCCAGCCAGCCAGCTTTCTCGGAAAA 180
Qy 181 CGCGGAAAAAGTGAAGCGCGCATGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGAAGCGCGCATGCGGAGCTGAATTAATTCCTCCAAACCGCGTGGCAC 240
Qy 241 AACAACTGGCGGGCAAAACAGTCTGTTGCTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGCAAAACAGTCTGTTGCTGATTTGGGTTGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGCGCGTCCCAATTCGCGCGCATTAATCTCGCGCGCATCAATCTCGCGCGCATCAATCTGGGTGCAGCG 360
Db 301 ACGCGCGTCCCAATTCGCGCGCATTAATCTCGCGCGCATCAATCTCGCGCGCATCAATCTGGGTGCAGCG 360
Qy 361 TGGTGGTTCGATGGTAGAACAGCGCGCTCGAAGCTCTGAAAGCGCGGTGCAACAATC 420
Db 361 TGGTGGTTCGATGGTAGAACAGCGCGCTCGAAGCTCTGAAAGCGCGGTGCAACAATC 420
Qy 421 TTCTCGCGCAACCGCGTCAAGTGGGTGATCATTAATCTATCCGTGGATGACCAAGATGCCA 480
Db 421 TTCTCGCGCAACCGCGTCAAGTGGGTGATCATTAATCTATCCGTGGATGACCAAGATGCCA 480
Qy 481 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCGGCGGTTATTTCTTGATCTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTGCCCTGCACTAATGTTCGGCGGTTATTTCTTGATCTCTGACCCAGA 540
Qy 541 CACCCATCAACAGATTAATTTTCTCCCATGAAGCGGTACGCGACTGGCGCGGTGAGAGCATC 600
Db 541 CACCCATCAACAGATTAATTTTCTCCCATGAAGCGGTACGCGACTGGCGCGGTGAGAGCATC 600
Qy 601 TGGTGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGGCCCATTAAGTTCTGTCTCGG 660

Db 601 |||||TGCTCGCATTTGGTCAACGAGCAAAATCGCGCTGTGTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
QY 721 CGGAACGGGAAGCGGACTGAGTGCATGTCCGGTTTTCAACAAACCATGCAATTCGCTGA 780
Db 721 CGGAACGGGAAGCGGACTGAGTGCATGTCCGGTTTTCAACAAACCATGCAATTCGCTGA 780
QY 781 ATGAGGGCATCGTCCCACTGGCATGCTGGTTGCCAAACGATCAGATGGCGTGGCGCAA 840
Db 781 ATGAGGGCATCGTCCCACTGGCATGCTGGTTGCCAAACGATCAGATGGCGTGGCGCAA 840
QY 841 TCGCGGCCATTACCGAGTCCGGGCTGGCGTTGGTGGCGATATCTCGTGTGGGATACG 900
Db 841 TCGCGGCCATTACCGAGTCCGGGCTGGCGTTGGTGGCGATATCTCGTGTGGGATACG 900
QY 901 AGGATACCGAAGACAGCTCATGTTATATCCGCCGTTTAAACCAATCAAAACAGGATTTTC 960
Db 901 AGGATACCGAAGACAGCTCATGTTATATCCGCCGTTTAAACCAATCAAAACAGGATTTTC 960
QY 961 GGCTGCTGGGCAACACGAGTGGACCGCTTCTCACTGGTGAAGAAACCAACCTCGCGGCCAATA 1020
Db 961 GGCTGCTGGGCAACACGAGTGGACCGCTTCTCACTGGTGAAGAAACCAACCTCGCGGCCAATA 1020
QY 1021 AGGCAATCAGCTGTGGCCGTTCTCACTGGTGAAGAAACCAACCTCGCGGCCAATA 1080
Db 1021 AGGCAATCAGCTGTGGCCGTTCTCACTGGTGAAGAAACCAACCTCGCGGCCAATA 1080
QY 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGTGGACGACGAGGTTT 1140
Db 1081 CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATGAGTGGACGACGAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGAGTGAGCGCAACGCAATTAATGAGTGGACGACGAGGTTT 1200
Db 1141 CCCGACTGGAAGCGGAGTGAGCGCAACGCAATTAATGAGTGGACGACGAGGTTT 1200
QY 1201 GCACAAATCTCATGTTGACAGCTTATCATCGACTGACCGTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTGACAGCTTATCATCGACTGACCGTGACCAATGCTTCTGGCG 1260
QY 1261 TCAGCAGCCATCGGAAGCTGTGATGCTGTGCGAGTGTGAATCACTGCAATAATTCG 1320
Db 1261 TCAGCAGCCATCGGAAGCTGTGATGCTGTGCGAGTGTGAATCACTGCAATAATTCG 1320
QY 1321 TGTGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCATACCGGTT 1380
Db 1321 TGTGCTCAAGGCGCACTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCATACCGGTT 1380
QY 1381 CTGGCAATATTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
Db 1381 CTGGCAATATTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA 1440
QY 1441 ATTGTGAGCGGATAACAATTTTACACAGGAACAGCCAGTCCGTTTAGGTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAATTTTACACAGGAACAGCCAGTCCGTTTAGGTTTTCACGA 1500
QY 1501 GCACTTCAACCAAGGACCATAGATTATGAATACTGAAGAGTAACTGGTAATCTGG 1560
Db 1501 GCACTTCAACCAAGGACCATAGATTATGAATACTGAAGAGTAACTGGTAATCTGG 1560
QY 1561 ATTTAACCGGATTAAGGCTATACCGTCTCGTGAAGTCCGTTAAGAAATTCGAGAAGAT 1620
Db 1561 ATTTAACCGGATTAAGGCTATACCGTCTCGTGAAGTCCGTTAAGAAATTCGAGAAGAT 1620
QY 1621 ACCGGAATTAAGTACCGTTGAGCTCCGATAAATCTGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTACCGTTGAGCTCCGATAAATCTGGAAGAGAAATTCACAGGTT 1680
QY 1681 CGCGCAATCGCGATGGCCCTGACATTAATCTTCTGGGCAACGACCGCTTGTGGCTAC 1740

Db 1681 CGCGCAATCGCGATGGCCCTGACATTAATCTTCTGGGCAACGACCGCTTGTGGGTAC 1740
QY 1741 GCTCAATCTGGCCTGTTGGCTGAATAATCAACCCCGCAAAAGGTTCCAGGCAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAATAATCAACCCCGCAAAAGGTTCCAGGCAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACGGCAAGCTGATTCGTTACCGGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTCAACGGCAAGCTGATTCGTTACCGGATCGCTGTT 1860
QY 1861 GAACGTTATCGCTGATTTTATAACAAAGATCTGCTGCGAAACCCGCAAAACCTCGGAA 1920
Db 1861 GAACGTTATCGCTGATTTTATAACAAAGATCTGCTGCGAAACCCGCAAAACCTCGGAA 1920
QY 1921 GAGATCCGGGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGTATTCATC 1980
Db 1921 GAGATCCGGGCTGGATAAAGAACTGAAAGCGAAAGGTAAAGCGCGCTGTATTCATC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTCGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTCGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAACCGCAAGTACGACATTTAAAGACGTGGGGTGGATTAACGCTGCGGGAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTTAAAGACGTGGGGTGGATTAACGCTGCGGGAAGCG 2100
QY 2101 GGTCTGACCTTCTGCTGACCTGATTTAAACCAACACATGATGCGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGCTGACCTGATTTAAACCAACACATGATGCGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTTAAAGGCGAAACAGCGATGACCATCAACGCCCGTGG 2220
QY 2221 GCATGCTCAACATCGACACCGCAAAAGTGAATTTATGGTGAACGTTATCGCGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACCGCAAAAGTGAATTTATGGTGAACGTTATCGCGACCTTC 2280
QY 2281 AAGGCTCAACATCGCAACCGTTGCGTGGTGGTGGCGGAGGTTTAAACGCCCGAGT 2340
Db 2281 AAGGCTCAACATCGCAACCGTTGCGTGGTGGTGGCGGAGGTTTAAACGCCCGAGT 2340
QY 2341 CCGAAACAAAGAGCTGCGCAAAAGATTCCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAAACAAAGAGCTGCGCAAAAGATTCCTCGAAAACTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTTAAAGACAAACCGCTGGGTGGCTGAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTTAAAGACAAACCGCTGGGTGGCTGAGCGCTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCAGTATTGCCGCCACCATGGAAAAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTGCCGCCACCATGGAAAAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTGCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTGCTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGGAAGACCGCGACACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGTCAGACTGTCGATGAAGCCCTGGAAGACCGCGACACTAATTCGAGCTCG 2640
QY 2641 AACAAACAAACAAATAACAACTCAACCAACCTCGGGATCGAGGGAAGGATTTTCAGAATTC 2700
Db 2641 AACAAACAAACAAATAACAACTCAACCAACCTCGGGATCGAGGGAAGGATTTTCAGAATTC 2700
QY 2701 GATCTCTTCTCTGTTGGCCCGCGGCGCTTCGAGCCCGGGGAAGAGCCCTATGCTTGT 2760
Db 2701 GATCTCTTCTCTGTTGGCCCGCGGCGCTTCGAGCCCGGGGAAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTGGTAAGTCTTTCAGCCAGAGCAGCTCCCTGGTGGCGCACCCAGCGTACCCAC 2820
Db 2761 CCGGAATGTGGTAAGTCTTTCAGCCAGAGCAGCTCCCTGGTGGCGCACCCAGCGTACCCAC 2820

```
QY 2821 ACGGTGAAAAACCGTATATAATGCCAGAGTGGCGGCAAAATCTTTTAGCCAGAGCAGCAGC 2880
Db 2821 ACGGTGAAAAACCGTATATAATGCCAGAGTGGCGGCAAAATCTTTTAGCCAGTCCAGCAAC 2880
QY 2881 CTGGTGGCCCATCAACGCACTACTGGCGAGAGCCATACAAATGTCCTCAAGATGTGGC 2940
Db 2881 CTGGTGGCCCATCAACGCACTACTGGCGAGAGCCATACAAATGTCCTCAAGATGTGGC 2940
QY 2941 AAGTCTTTTCAGTGATGTGCGATCTTGGAGGCAACCAACGTAATCTCACACCGGGGAGAAG 3000
Db 2941 AAGTCTTTTCAGTGATGTGCGATCTTGGAGGCAACCAACGTAATCTCACACCGGGGAGAAG 3000
QY 3001 CCCTATGCTTGGCCGAATGTGTAAGTCTTCTCTCAGAGCTCTCACTGCTGGTGGCCAC 3060
Db 3001 CCCTATGCTTGGCCGAATGTGTAAGTCTTCTCAGAGCTCTCACTGCTGGTGGCCAC 3060
QY 3061 CAGCGTACCCACACCGGTGAAAAACCGTATATAATGCCAGAGTGGCGGCAAAATCTTTTAGC 3120
Db 3061 CAGCGTACCCACACCGGTGAAAAACCGTATATAATGCCAGAGTGGCGGCAAAATCTTTTAGC 3120
QY 3121 CGCAGCGATAACCTGGTGGCGCATCAACGCACTACTACTGGCGAGAGCCATACAAATGT 3180
Db 3121 CAGCGCGCCCACTTGAACGCCATCAACGCACTACTACTGGCGAGAGCCATACAAATGT 3180
QY 3181 CCAGATGTGGCAAGTCTTTCTCAACTTCAGGCCAATTTGGTCCGTCAACAGTACTCAC 3240
Db 3181 CCAGATGTGGCAAGTCTTTCTCAACTTCAGGCCAATTTGGTCCGTCAACAGTACTCAC 3240
QY 3241 ACCGCTAAAAAATAGTGGCCAGCGCGCCAGTACCGGTACGAGCTTCGCGACTACGCT 3300
Db 3241 ACCGCTAAAAAATAGTGGCCAGCGCGCCAGTACCGGTACGAGCTTCGCGACTACGCT 3300
```

RESULT 3

```
US-09-765-555-14
; Sequence 14, Application US/0976555
; Publication No. US20030037355A1
; GENERAL INFORMATION:
; APPLICANT: The Scripps Research Institute
; TITLE OF INVENTION: Methods and compositions to modulate
; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US/09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m1 and zinc finger
; OTHER INFORMATION: protein 2FPm1
US-09-765-555-14
```

```
Query Match 97.4%; Score 3215.2; DB 10; Length 3300;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 3247; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 1 CCGACACCATCAATGGTGCAGAAACCTTTCGGGTATGCGATGATAGCGCCGGAGAGA 60
Db 1 CCGACACCATCAATGGTGCAGAAACCTTTCGGGTATGCGATGATAGCGCCGGAGAGA 60
QY 61 GTCAATTTCAGGTGGTGAATGTGAACACAGTAAACGTTATACGATGTGCGAGAGTATGCG 120
Db 61 GTCAATTTCAGGTGGTGAATGTGAACACAGTAAACGTTATACGATGTGCGAGAGTATGCG 120
QY 121 GTGTCCTTTATCAGACCGGTTTCCCGGTGGTGAACACGAGCCACCGCTTCTCGCAAAA 180
```

```
Db 121 GTGTCCTTTATCAGACCGGTTTCCCGGTGGTGAACACGAGCCAGCCACGCTTCTCGCAAAA 180
QY 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTAATATCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGGAAAGCGCGATGGCGAGCTGAATTAATATCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATGTGGGTTGCCACCTCAGTCAGTCGCGCCCTGC 300
Db 241 AACAACTGGCGGGCAAAACAGTCGTTGCTGATGTGGGTTGCCACCTCAGTCAGTCGCGCCCTGC 300
QY 301 ACGCGCGTCCCAAAATGTTCGCGCGATTAATATCTCGCGCGATCAACCTGGGTGGCCACG 360
Db 301 ACGCGCGTCCCAAAATGTTCGCGCGATTAATATCTCGCGCGATCAACCTGGGTGGCCACG 360
QY 361 TGGTGGTGTGATGTAAGCAAGCGCGCTGGAAGCGCTGTAAGCGCGGTGACACAATC 420
Db 361 TGGTGGTGTGATGTAAGCAAGCGCGCTGGAAGCGCTGTAAGCGCGGTGACACAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATATCTCCGCTGGATGACCAAGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATTAATATCTCCGCTGGATGACCAAGATGCCA 480
QY 481 TTGCTGTGGAGAGCTGCCCTGCACTAATGTTCGCGGCTTAATTTCTTGATGCTCTGACCCAGA 540
Db 481 TTGCTGTGGAGAGCTGCCCTGCACTAATGTTCGCGGCTTAATTTCTTGATGCTCTGACCCAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
QY 601 TGGTCGCATTTGGGTACACAGCAAAATCGCGCTGTGTAGCGGGCCCATTAAGTCTGTCTCGG 660
Db 601 TGGTCGCATTTGGGTACACAGCAAAATCGCGCTGTGTAGCGGGCCCATTAAGTCTGTCTCGG 660
QY 661 CGCGTCTCGCTCGCTGGCTGGCATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTCGCTCGCTGGCTGGCATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGGGACCTGGAAGTGCATGTCGGGTTTTCACAAACCACTGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGGACCTGGAAGTGCATGTCGGGTTTTCACAAACCACTGCAAAATGCTGA 780
QY 781 ATGAGGGCATCGTTCCCACTCGCATGTGTTGTCACACGATCAGATGCGCTGGCGCGCAA 840
Db 781 ATGAGGGCATCGTTCCCACTCGCATGTGTTGTCACACGATCAGATGCGCTGGCGCGCAA 840
QY 841 TGGCGGCATTTACGAGTCCGGGCTGGCGTGGTGGCGATATCTCGTGTAGTGGATAGC 900
Db 841 TGGCGGCATTTACGAGTCCGGGCTGGCGTGGTGGCGATATCTCGTGTAGTGGATAGC 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCTTAAACCACTCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGCTTAAACCACTCAAAACAGGATTTTC 960
QY 961 GCCTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGTGTGGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTGCCCCCTCTCACTGGTGAAGAAAAAACCCCTTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTGCCCCCTCTCACTGGTGAAGAAAAAACCCCTTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAATGAGCTGGCACGACAGGTTT 1140
Db 1081 CGCAAAACCGCTCTCTCCCGCGCTTGGCGGATTCATTAATGAGCTGGCACGACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGAGTGTAGCTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGGCAGTGAGCGCAACGCAATTAATGAGTGTAGCTCACTCATTTAG 1200
QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGGTGCAACCAATGCTTCTGGGG 1260
Db 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCGGTGCAACCAATGCTTCTGGGG 1260
```


Db	1201	GCACAAATTTCTCATGTTTGACAGCTTATCATCGATGCAAGGTGCACCAATGCTTCTCGCG	1261
Qy	1261	TCAGGAGCCATCGGAAGCTGTGGTATCGCTGTGCAGGTCGTAATAACACTGCATAAATCG	1320
Db	1261	TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGCAGGTCGTAATAACACTGCATAAATCG	1320
Qy	1321	TGTCGCTCAAGCGCACATCCCGTCTCGATAAATGTTTTTTCGGCGGACATCATCAACGGTT	1380
Db	1321	TGTCGCTCAAGCGCACATCCCGTCTCGATAAATGTTTTTTCGGCGGACATCATCAACGGTT	1380
Qy	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA	1440
Db	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATAATGTGTGA	1440
Qy	1441	ATTGTGAGCGGATAACAATTTTACACAGGAAACGCCAGTCCGTTTATAGGTGTTTTACGA	1500
Db	1441	ATTGTGAGCGGATAACAATTTTACACAGGAAACGCCAGTCCGTTTATAGGTGTTTTACGA	1500
Qy	1501	GCATTTACCAACAAAGGACCATAGATTATGAAACTGAAGAAGTGAACCTGGTAACTCTGG	1560
Db	1501	GCATTTACCAACAAAGGACCATAGATTATGAAACTGAAGAAGTGAACCTGGTAACTCTGG	1560
Qy	1561	ATTAAACGCGGATAAAGGCTATAACGGTCTCGCTGAACTCGGTAAAGAAATTCGAGAAAGAT	1620
Db	1561	ATTAAACGCGGATAAAGGCTATAACGGTCTCGCTGAACTCGGTAAAGAAATTCGAGAAAGAT	1620
Qy	1621	ACCGGAATTTAAGTCAACGTTGAGCATCCGGATAAATCGGAAGAGTAACCTGGTAACTCTGG	1680
Db	1621	ACCGGAATTTAAGTCAACGTTGAGCATCCGGATAAATCGGAAGAGTAACCTGGTAACTCTGG	1680
Qy	1681	CGGCAACTGGCGATGCGCCTGACATATTCTTTGGGCACACGACCGCTTTGGTGGCTAC	1740
Db	1681	CGGCAACTGGCGATGCGCCTGACATATTCTTTGGGCACACGACCGCTTTGGTGGCTAC	1740
Qy	1741	GCTCAATCTGGCCTGTTGGCTGGAATCACCCCGACAAAGGTTTCCAGGACAAGCTGTAT	1800
Db	1741	GCTCAATCTGGCCTGTTGGCTGGAATCACCCCGACAAAGGTTTCCAGGACAAGCTGTAT	1800
Qy	1801	CCGTTTACCTGGGATGCCGTACGTTACAAACGGCAAGCTGATTGCTTACCCCGATCGCTGTT	1860
Db	1801	CCGTTTACCTGGGATGCCGTACGTTACAAACGGCAAGCTGATTGCTTACCCCGATCGCTGTT	1860
Qy	1861	GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCCGAACCGCCAAACAACTGGGAA	1920
Db	1861	GAAGCGTTATCGCTGATTATTAACAAGATCTGCTGCCGAACCGCCAAACAACTGGGAA	1920
Qy	1921	GAGATCCCGCGCTGGATAAAGAACTCAAAAGCGAAAGGTAAAGACGCGCTGATGTTCAAC	1980
Db	1921	GAGATCCCGCGCTGGATAAAGAACTCAAAAGCGAAAGGTAAAGACGCGCTGATGTTCAAC	1980
Qy	1981	CTGCAAGAACGTTACTTCACTGCGCCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG	2040
Db	1981	CTGCAAGAACGTTACTTCACTGCGCCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG	2040
Qy	2041	TATGAAAAACGCAAGTACGACATTTAAAGACGTGGGCTGATTAACGCTGCGCGGAAGCG	2100
Db	2041	TATGAAAAACGCAAGTACGACATTTAAAGACGTGGGCTGATTAACGCTGCGCGGAAGCG	2100
Qy	2101	GGTCTGACCTTCTCGTTTGACCTGATTTAAAAACAAACACATGAATCGACACCGGATTAC	2160
Db	2101	GGTCTGACCTTCTCGTTTGACCTGATTTAAAAACAAACACATGAATCGACACCGGATTAC	2160
Qy	2161	TCCATCCGAGAAGCTGCTTTTAAAGGCGAAACACGATGACCATCAACGGCCCGTGG	2220
Db	2161	TCCATCCGAGAAGCTGCTTTTAAAGGCGAAACACGATGACCATCAACGGCCCGTGG	2220
Qy	2221	GCATGFTCCAAATCGACACCGACAAAGTGAATTTATGGTGTAAACGGTACTGGCGACCTTC	2280
Db	2221	GCATGFTCCAAATCGACACCGACAAAGTGAATTTATGGTGTAAACGGTACTGGCGACCTTC	2280
Qy	2281	AAGGGTCAACATCCAAACCGTTCTGTTGGCGTCTGAGCGCAGGTATTAAACGGCCAGT	2340
Db	2281	AAGGGTCAACATCCAAACCGTTCTGTTGGCGTCTGAGCGCAGGTATTAAACGGCCAGT	2340

RESULT 4

US-09-765-555-15

US-05-703-333-13
; Sequence 15, Application US/09765555

; Sequence ID, Application US/03
; Publication No. US20030037355A1

; PUBLICATION NO. US20
; GENERAL INFORMATION:
; GENERAL INFORMATION:

APPLICANT: The Scripps Research Institute

; TITLE OF INVENTION: Methods and compositions to modulate

```
; TITLE OF INVENTION: expression in plants
; FILE REFERENCE: 27801-20014.40
; CURRENT APPLICATION NUMBER: US 09/765,555
; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: US 09/620,897
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US 60/177,468
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 3300
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Partial sequence of pMal-m2 and zinc finger
; OTHER INFORMATION: protein ZFPm2
US-09-765-555-15

Query Match          97.4%; Score 3215.2; DB 10; Length 3300;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 3247; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 1 CGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
DB 1 CGACACCATCGAATGGTGCAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTCAGGGTGGTGAATGTGAACACCAAGTAACTTATACGATGTCGACAGTATGCCG 120
DB 61 GTCAATTCAGGGTGGTGAATGTGAACACCAAGTAACTTATACGATGTCGACAGTATGCCG 120
QY 121 GTGTCTTATCAGACCGTTTCCCGGTGGTGAACACCAAGTAACTTATACGATGTCGACAGT 180
DB 121 GTGTCTTATCAGACCGTTTCCCGGTGGTGAACACCAAGTAACTTATACGATGTCGACAGT 180
QY 181 CGCGGAAAAAGTGAAGCGGGGATGGCGAGCTGAATTTACATTTCCCAACCGCTGGGCAC 240
DB 181 CGCGGAAAAAGTGAAGCGGGGATGGCGAGCTGAATTTACATTTCCCAACCGCTGGGCAC 240
QY 241 AACAACTGGCGGCAACAGTGGTGTCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGCAACAGTGGTGTCTGATTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
QY 301 AGCGCCGTGCGAAATTTGTCGGCGGATTAATCTCGGCGGATCAACTGGGTGCGACGG 360
DB 301 AGCGCCGTGCGAAATTTGTCGGCGGATTAATCTCGGCGGATCAACTGGGTGCGACGG 360
QY 361 TGGTGGTGTGATGTAGAACGAAGCGGCTGGAAGCCTGTAAAGCGGCGGTGCAATC 420
DB 361 TGGTGGTGTGATGTAGAACGAAGCGGCTGGAAGCCTGTAAAGCGGCGGTGCAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGGCTGGATGACCAAGGATGCCA 480
DB 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTCGGCTGGATGACCAAGGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCCTGCATTAATGTTCCGGCGTATTTCTGATGCTCTGACACAGA 540
DB 481 TTGCTGTGGAAGCTGCCTGCATTAATGTTCCGGCGTATTTCTGATGCTCTGACACAGA 540
QY 541 CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC 600
DB 541 CACCCATCAACAGTATTTTCTCCCATGAAGACGGTACCGGACTGGGCGTGGAGCATC 600
QY 601 TGGTGGCATTTGGTCAACAGAAATCGCGCTGTAGCGGGCCCAATTAAGTCTCTGTCGG 660
DB 601 TGGTGGCATTTGGTCAACAGAAATCGCGCTGTAGCGGGCCCAATTAAGTCTCTGTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
DB 661 CGCGTCTGCGTCTGGCTGGCATTAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGCGCATGGAGTGCATGTCGGTTCACAAACCAACCATGCAAAATGCTGA 780
DB 721 CGGAACGGGAAGCGCATGGAGTGCATGTCGGTTCACAAACCAACCATGCAAAATGCTGA 780
```

```
DB 721 CGGAACGGGAAGCGCATGGAGTGCATGTCGGTTCACAAACCAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATCTGTTCCCATCGCATGTGGTTGCGCAACGATCAGATGCGCTGGCGCGAA 840
DB 781 ATGAGGGCATCTGTTCCCATCGCATGTGGTTGCGCAACGATCAGATGCGCTGGCGCGAA 840
QY 841 TGGCGCCATTACCGAGTCCGGGCTGGCGGTGGTGGCGGATATCTCGGTAGTGGGATAG 900
DB 841 TGGCGCCATTACCGAGTCCGGGCTGGCGGTGGTGGCGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTTATATCCCGCGCTTAACCAACCATCAAAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTTATATCCCGCGCTTAACCAACCATCAAAACAGGATTTTC 960
QY 961 GCCTGTGGGGCAACCAACGAGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
DB 961 GCCTGTGGGGCAACCAACGAGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTTGCCCTCTCACTGTGTAAGAAAGAAACCAACCTGGCGCCCAATA 1080
DB 1021 AGGCAATCAGCTGTTGCCCTCTCACTGTGTAAGAAAGAAACCAACCTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAAATGAGCTGGCAGACACAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGCTTGGCGGATTCATTAAATGAGCTGGCAGACACAGGTTT 1140
QY 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTCAGTCACTCAATAG 1200
DB 1141 CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTAGTTCAGTCACTCAATAG 1200
QY 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGGTGCACCAATGCTTCTGGCG 1260
DB 1201 GCACAAATCTCATGTTTGACAGCTTATCATCGACTGCAAGGTGCACCAATGCTTCTGGCG 1260
QY 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAATCACTGCATAAATCG 1320
DB 1261 TCAGCAGCCATCGGAAGCTGTGGTATGGCTGTGAGGTGCGTAATCACTGCATAAATCG 1320
QY 1321 TGTGCTCAAGCGGCACTCCCGTCTGGATAATGTTTTTGGCGCGACATCAATACGGTT 1380
DB 1321 TGTGCTCAAGCGGCACTCCCGTCTGGATAATGTTTTTGGCGCGACATCAATACGGTT 1380
QY 1381 CTGGCAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGTGA 1440
DB 1381 CTGGCAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGTGA 1440
QY 1441 ATTGTAGCGGATAAACAAATTTCAACAGAAACAGCCAGTCCGTTTAGGTGTTTACGA 1500
DB 1441 ATTGTAGCGGATAAACAAATTTCAACAGAAACAGCCAGTCCGTTTAGGTGTTTACGA 1500
QY 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAGGTAAACTGGTAACTCTGG 1560
DB 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACCTGAAGAGGTAAACTGGTAACTCTGG 1560
QY 1561 ATTAACCGCGATTAAGGCTATAAACGCTCTCGTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTAACCGCGATTAAGGCTATAAACGCTCTCGTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACACAGGTT 1680
DB 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACACAGGTT 1680
QY 1681 GCGGCAACTGGCGATGGCCCTGACATATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
DB 1681 GCGGCAACTGGCGATGGCCCTGACATATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCGGAGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
DB 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCGGAGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGATGCGGTACGTTACAAACCGGAGCTGATTTGCTTACCCGATCGCTGT 1860
DB 1801 CCGTTTACCTGGATGCGGTACGTTTACAAACCGGAGCTGATTTGCTTACCCGATCGCTGT 1860
```


Db 241 AACAACTGGCGGCAAAACAGTCTGTTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
Qy 301 ACGCGCCGTGCGAAATTTGTGGCGCGATTAATCTCGCGCCGATCAATCTGGGTGCCAGC 360
Db 301 ACGCGCCGTGCGAAATTTGTGGCGCGATTAATCTCGCGCCGATCAATCTGGGTGCCAGC 360
Qy 361 TGGTGGTCTCGATGTAGAACGAAAGCGCGTGAAGCCTGTAAAGCGCGGTGCAACAATC 420
Db 361 TGGTGGTCTCGATGTAGAACGAAAGCGCGTGAAGCCTGTAAAGCGCGGTGCAACAATC 420
Qy 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACCAAGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGCTGATCAATTAATCTATCCGCTGGATGACCAAGATGCCA 480
Qy 481 TTGCTGTGGAAGCTCGCTGCAATATGTTCCGGCGTATTTCTTGATGTCTCTGACCCAGA 540
Db 481 TTGCTGTGGAAGCTCGCTGCAATATGTTCCGGCGTATTTCTTGATGTCTCTGACCCAGA 540
Qy 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Qy 601 TGGTGGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Db 601 TGGTGGCATTTGGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCAATTAAGTTCTGTCTCGG 660
Qy 661 CGCGTCTGGCTGCGTGGCTGCGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGGCTGCGTGGCTGCGCAATAAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTGTCCCAACGATCAGATGGCGCTGGCGCGAA 780
Db 721 CGGAACGGGAAGGCACTGGAGTGCATGTCGGTGTCCCAACGATCAGATGGCGCTGGCGCGAA 780
Qy 781 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTGTCCCAACGATCAGATGGCGCTGGCGCGAA 840
Db 781 ATGAGGGCATCGTTCCTCACTGCGATGCTGGTGTCCCAACGATCAGATGGCGCTGGCGCGAA 840
Qy 841 TGGCGCCATTACCGAGTCCGGGCTGCGGCTGGTGGTGGGATATCTCGGTAGTGGGATAGC 900
Db 841 TGGCGCCATTACCGAGTCCGGGCTGCGGCTGGTGGTGGGATATCTCGGTAGTGGGATAGC 900
Qy 901 ACGATACGGAAGACGCTCATGTTATATCCCGCCGTTAAACCAACATCAACAGGATTTTC 960
Db 901 ACGATACGGAAGACGCTCATGTTATATCCCGCCGTTAAACCAACATCAACAGGATTTTC 960
Qy 961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAAAACAGCGTGGACCGCTTCTGCAACTCTCTCAGGGCCAGGCGGTGA 1020
Qy 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAAGAAACCAACCTGGCGCCCAATA 1080
Qy 1081 CGCAAAACGCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAAGAGTTTC 1140
Db 1081 CGCAAAACGCTCTCTCCCGCGGTTGGCGGATTCATTAATGAGCTGGCAAGAGTTTC 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGCAGCGCAACCAATTAATGAGTTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGCAGCGCAACCAATTAATGAGTTAGTCACTCATTTAG 1200
Qy 1201 GCACAATTTCTCATGTTTGAAGCTTATCATGACCTGCAGGTGCAACCAATGCTTCTGGCG 1260
Db 1201 GCACAATTTCTCATGTTTGAAGCTTATCATGACCTGCAGGTGCAACCAATGCTTCTGGCG 1260
Qy 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGAGAGTTCGTAATTCATGTCATTAATTCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGGTATGCTGTGAGAGTTCGTAATTCATGTCATTAATTCG 1320
Qy 1321 TGTTCCTCAAGCGCACTCCCGTTCGTGATATGTTTTTGGCGGACATCATACGGTT 1380
Db 1321 TGTTCCTCAAGCGCACTCCCGTTCGTGATATGTTTTTGGCGGACATCATACGGTT 1380

Qy 1381 CTGGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTGTGA 1440
Db 1381 CTGGCAATATTTCTGAATGAGCTGTTGACAAATTAATCATCGGCTCGTATTAATGTGTGA 1440
Qy 1441 ATTGTGAGCGGATAAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Db 1441 ATTGTGAGCGGATAAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Qy 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAATCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAATCTGG 1560
Qy 1561 ATTAACGCGGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGCGGATAAAGGCTATAACCGTCTCGCTGAAGTCGGTAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTTAAAGTCACCGTTGAGCATCCGATATAACTGGAAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTTAAAGTCACCGTTGAGCATCCGATATAACTGGAAAGAGAAATTCACAGGTT 1680
Qy 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACGACCCGCTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCAACGACCCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCGGAACAAAGGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCGGAACAAAGGTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGATGGCGTACGTTTACAAACGCAAGCTGATTGTTACCCGATCCCTGTT 1860
Db 1801 CCGTTTACCTGGATGGCGTACGTTTACAAACGCAAGCTGATTGTTACCCGATCCCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAAAACCCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCGAACCCGCCAAAAAACCCTGGGAA 1920
Qy 1921 GAGATCCCGCGCTGGATATAAGAACTGAAACGGAAGGTAAAGAGCGCGTGTATTCAC 1980
Db 1921 GAGATCCCGCGCTGGATATAAGAACTGAAACGGAAGGTAAAGAGCGCGTGTATTCAC 1980
Qy 1981 CTGCAAGAACCGTACTTCCACCTGGCGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCCACCTGGCGCTGATGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAAAACGCAAGTACGACATTAAGAACGCTGGCGCTGGATAAACGCTGGCGCAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGAACGCTGGCGCTGGATAAACGCTGGCGCAAGCG 2100
Qy 2101 GGTCTGACCTTCTCGTGGTGAACGATTAATAAACAACAACATGAATGCAAGACCCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTGGTGAACGATTAATAAACAACAACAACATGAATGCAAGACCCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTGCCCTTAAATAAGGCGAACAAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTGCCCTTAAATAAGGCGAACAAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGCTCAACATCGACACCAAGTGAATTAATGTTAAACCGTACTGCGGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACCAAGTGAATTAATGTTAAACCGTACTGCGGACCTTC 2280
Qy 2281 AAGGTCACCATCCAAACCGTTGCTGGCGTGGTGGCGGATGATTAAGCGGCCAGT 2340
Db 2281 AAGGTCACCATCCAAACCGTTGCTGGCGTGGTGGCGGATGATTAAGCGGCCAGT 2340
Qy 2341 CGGAACAGAGCTGGGAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGCTG 2400
Db 2341 CGGAACAGAGCTGGGAAAAGAGTTCCTCGAAAACTATCTGCTGACTGATGAAGGCTG 2400
Qy 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGTGAAGTCTTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGTGAAGTCTTTACGAGGAAG 2460

```
QY 2461 TTGGCGAAGATCCAGCTATTGTCGCCACCATGGAAGAGGCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAGATCCAGCTATTGTCGCCACCATGGAAGAGGCGCCAGAAAGGTGAATCATG 2520
QY 2521 CGGAACATCCGCGAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAAGCC 2580
Db 2521 CGGAACATCCGCGAGATGTCGCTTTCTGGTATGCGGTGCGTACTGCGGTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGTCAGACTGTGATGAAGCCCTTGAAGAGCGGCGAGACTAAATTCAGAGCTCG 2640
Db 2581 GCCAGCGTCTGTCAGACTGTGATGAAGCCCTTGAAGAGCGGCGAGACTAAATTCAGAGCTCG 2640
QY 2641 AACACACACACATTAACAATAAACAACCTTCGGGATCGAGGAGGAGATTCAGAAATTC 2700
Db 2641 AACACACACACATTAACAATAAACAACCTTCGGGATCGAGGAGGAGATTCAGAAATTC 2700
QY 2701 GGATCCTCTTCTGTCGCGCCAGCGGCGCTTCGAGCGCGGCGGAGAGCCCTATGCTTGT 2760
Db 2701 GGATCCTCTTCTGTCGCGCCAGCGGCGCTTCGAGCGCGGCGGAGAGCCCTATGCTTGT 2760
QY 2761 CCGGAATGTGTAAGTCTTTCAGCCAGAGAGCTCCCTGCTGCGCCACAGCGCTACCCAC 2820
Db 2761 CCGGAATGTGTAAGTCTTTCAGCCAGATCCTGCGCCACCTGCTGCGCCACAGCGTACCCAC 2820
QY 2821 ACGGTTGAAAACCGTATAAATGCCAGAGTGGGCAAAATCTTTAGCCAGAGCAGCAGC 2880
Db 2821 ACGGTTGAAAACCGTATAAATGCCAGAGTGGGCAAAATCTTTAGCCAGCAGCGGCTCC 2880
QY 2881 CTGTCGCGCCATCAACGACTCATACTGGCGAGAGCCATACAAATGTCAGAGATGTGGC 2940
Db 2881 CTGTCGCGCCATCAACGACTCATACTGGCGAGAGCCATACAAATGTCAGAGATGTGGC 2940
QY 2941 AAGTCTTTCACTGATGTCGTGATCTTTCGAGGAGCACCACGCTACTCACCCGGGAGAG 3000
Db 2941 AAGTCTTTCACTGAGAGCTCCAGCGCTGCTGCGCCACCAACGCTACTCACCCGGGAGAG 3000
QY 3001 CCCTATGCTTTCGCGGATGTTGTAAGTCTTCTCAGAGCTCTCAGAGCTCTCAGCTGCGCCAC 3060
Db 3001 CCCTATGCTTTCGCGGATGTTGTAAGTCTTCTCAGAGCAGAGCTCTCAGCTGCGCCAC 3060
QY 3061 CAGGTTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTAGC 3120
Db 3061 CAGGTTACCCACACGGGTGAAAAACCGTATAAATGCCAGAGTGGCGCAAAATCTTTAGT 3120
QY 3121 CGAGCGATTAACCTGTCGCGCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
Db 3121 GACTGCGCGCACTTGTCTGCGCATCAACGCACTCATCTGCGGAGAGCCATACAAATGT 3180
QY 3181 CCAGAAATGTGCAAGTCTTTCTCAACTTCAGGCGCATTTGGTCCGTCACCACTACTCAC 3240
Db 3181 CCAGAAATGTGCAAGTCTTTCTCAACTTCAGGCGCATCTCGTCGCGCACCAACGCTACTCAC 3240
QY 3241 ACCGGTAAAAAACTAGTGGCGAGCGCGGCGGCGAGTACCGCTACGACGTTCCGGACTACGCT 3300
Db 3241 ACCGGTAAAAAACTAGTGGCGAGCGCGGCGGCGGAGTACCGCTACGACGTTCCGGACTACGCT 3300

RESULT 6
US-10-257-384A-1
; Sequence 1, Application US/10257384A
; Publication No. US20040087524A1
; GENERAL INFORMATION:
; APPLICANT: Wiederanders, Bernd
; APPLICANT: Maubach, Gunter
; TITLE OF INVENTION: Agent for postoperative use after removal of bone tumors
; FILE REFERENCE: 2945-101
; CURRENT APPLICATION NUMBER: US/10/257,384A
; PRIORITY FILING DATE: 2003-08-21
; PRIORITY FILING DATE: PCT/DE 01/01510
; PRIORITY FILING DATE: 2001-04-18
; PRIORITY FILING DATE: 2000-04-18
; NUMBER OF SEQ ID NOS: 19
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 7373
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Expression vector pMalc2chbnp2
; NAME/KEY: CDS
; LOCATION: (1528)..(3453)
; OTHER INFORMATION: /product = fusion protein MBP Cystatin C BMP-2
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2689)
; OTHER INFORMATION: Faktor Xa Protease /position=388
; OTHER INFORMATION: (Amino acid sequence)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3067)..(3108)
; OTHER INFORMATION: /spacer between Cystatin C andBMP-2
; PUBLICATION INFORMATION:
; AUTHORS: Zwick, M B.
; TITLE: Expression vector pMal-X, complete sequence
; JOURNAL: Anal. Biochem.
; VOLUME: 264
; ISSUE: 1
; PAGES: 87-97
; DATE: 1998-11-01
; DATABASE ACCESSION NUMBER: Entrez Nucleotide database/AF031813
; DATABASE ENTRY DATE: 2001-05-07
; RELEVANT RESIDUES: 1528 - 3453
US-10-257-384A-1
```

```
Query Match 82.0%; Score 2705.8; DB 18; Length 7373;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 2719; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTCGCGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCAATTCAGGTTGGTGAATGTGAACACAGTAACTGTTATACGATGTGCGAGAGTATGCCG 120
Db 61 GTCAATTCAGGTTGGTGAATGTGAACACAGTAACTGTTATACGATGTGCGAGAGTATGCCG 120
QY 121 GTGCTCTTATCAGACCGTTTCCCGGTGTGTAACAGGCGGCGGCGGCGGCGGCGGCGG 180
Db 121 GTGCTCTTATCAGACCGTTTCCCGGTGTGTAACAGGCGGCGGCGGCGGCGGCGGCGG 180
QY 181 CGCGGAAAAAGTGGAAAGCGGCGGATGCGGAGCTGAATTAATTCACATTCGCAACCGCGTGGC 240
Db 181 CGCGGAAAAAGTGGAAAGCGGCGGATGCGGAGCTGAATTAATTCACATTCGCAACCGCGTGGC 240
QY 241 AACCACTGGCGGCGGCAACAGTCTGCTGATGTTGGCGGTGCGCACTCAGTCTGGCCCTGC 300
Db 241 AACCACTGGCGGCGGCAACAGTCTGCTGATGTTGGCGGTGCGCACTCAGTCTGGCCCTGC 300
QY 301 ACGCGCGCTCGCAAAATTTGTCGGGCGGATTAATAATTCGCGCGGATCAACTGGGTGGCAGCG 360
Db 301 ACGCGCGCTCGCAAAATTTGTCGGGCGGATTAATAATTCGCGCGGATCAACTGGGTGGCAGCG 360
QY 361 TGGTGGTGTGATGTTAGAAAGCGGCGGCTGGAAGCTGTAAAGCGGCGGCGGTCACAAATC 420
Db 361 TGGTGGTGTGATGTTAGAAAGCGGCGGCTGGAAGCTGTAAAGCGGCGGTCACAAATC 420
QY 421 TTCTCGCGCAACCGGTCAGTGGGTGATCAATTAATTCGCTGGATGACGAGATGCCA 480
Db 421 TTCTCGCGCAACCGGTCAGTGGGTGATCAATTAATTCGCTGGATGACGAGATGCCA 480
QY 481 TTGCTGTGGAAGCTGCTGCACTTAATGTTCCGCGGCTTATTTCTTGATGTCTCTGACAGA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTTAATGTTCCGCGGCTTATTTCTTGATGTCTCTGACAGA 540
```

QY 541 CACCCATCAACAGTATTATTTCTCCATGAAGCGGTACGGACTGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTATTTCTCCATGAAGCGGTACGGACTGGCGTGGAGCATC 600
QY 601 TGGTCCGATTGGGTCAACAGCAAAATCGGCTGTAGCGGGCCCAATTAAGTCTCTCGG 660
Db 601 TGGTCCGATTGGGTCAACAGCAAAATCGGCTGTAGCGGGCCCAATTAAGTCTCTCGG 660
QY 661 CGCGTCTGGTCTGGCTGGCTGGCTGAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGGTCTGGCTGGCTGGCTGAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGCGGACTGGAGTGCATGTCGGTTCCTCAACAAACCAATGCAATGCTGA 780
Db 721 CGGAACGGGAAGCGGACTGGAGTGCATGTCGGTTCCTCAACAAACCAATGCAATGCTGA 780
QY 781 ATGAGGGCATCTTCCCACTGCGATGCTGTTGCCAAACGATCAGATGGCGCTGGCGCA 840
Db 781 ATGAGGGCATCTTCCCACTGCGATGCTGTTGCCAAACGATCAGATGGCGCTGGCGCA 840
QY 841 TGGCGCCATTACGAGTCCGGCTGGCGTGGTGGTGGGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGCCATTACGAGTCCGGCTGGCGTGGTGGTGGGATATCTCGGTAGTGGGATACG 900
QY 901 ACGATACGGAACAGCTCATGTTATATCCCGCGTTAAACCAATCAACAGGATTTTC 960
Db 901 ACGATACGGAACAGCTCATGTTATATCCCGCGTTAAACCAATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGCAACACGAGTGGACCGCTGCTGCAACTCTCAGGSCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGTGGACCGCTGCTGCAACTCTCAGGSCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCCGTCTCACTGGTGAAGAAACCAACCCCTGGCGCCCAATA 1080
QY 1081 CGCAACCGCCCTCTCCCGCGGTTGGCCGATTTCAATTAATGAGTGGGCAACAGAGTTT 1140
Db 1081 CGCAACCGCCCTCTCCCGCGGTTGGCCGATTTCAATTAATGAGTGGGCAACAGAGTTT 1140
QY 1141 CCCGACTGGAAGCGGAGTGGGCAACGAAATTAATGAGTGGTCACTCATTAG 1200
Db 1141 CCCGACTGGAAGCGGAGTGGGCAACGAAATTAATGAGTGGTCACTCATTAG 1200
QY 1201 GCACAAATCTCATGTTGACGCTTATCATCGACTGCACGGTGCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTGACGCTTATCATCGACTGCACGGTGCAATGCTTCTGGCG 1260
QY 1261 TCAGGAGCCATCGGAAGCTGTGATGCTGTGCAAGTTCGTAATCATCTGCATTAATCG 1320
Db 1261 TCAGGAGCCATCGGAAGCTGTGATGCTGTGCAAGTTCGTAATCATCTGCATTAATCG 1320
QY 1321 TGTGCTCAAGGCGCACTCCCGTCTGGGATAATGTTTTTGGCGGACATCAACGGTT 1380
Db 1321 TGTGCTCAAGGCGCACTCCCGTCTGGGATAATGTTTTTGGCGGACATCAACGGTT 1380
QY 1381 CTGGCAAAATATCTGAAAATGAGCTGTGACAATTAATCATCGGCTCGTATAATGTTGGA 1440
Db 1381 CTGGCAAAATATCTGAAAATGAGCTGTGACAATTAATCATCGGCTCGTATAATGTTGGA 1440
QY 1441 ATTGTAGCGGATAACAAATTTACAGAAACAGCCAGTCCGTTAGGTGTTTTCAACA 1500
Db 1441 ATTGTAGCGGATAACAAATTTACAGAAACAGCCAGTCCGTTAGGTGTTTTCAACA 1500
QY 1501 GCACCTTACCAACAGGACATAGATTATGAATCTGAAGAGGTTAACTGGTAACTCTCG 1560
Db 1501 GCACCTTACCAACAGGACATAGATTATGAATCTGAAGAGGTTAACTGGTAACTCTCG 1560
QY 1561 ATTTACCGGATAAAGGCTATAACGGTCTCGCTGAAGTCCGTTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTTACCGGATAAAGGCTATAACGGTCTCGCTGAAGTCCGTTAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGGATAAATCTGGAAGAGAAATTTCCACAGGTT 1680

Db 1621 ACCGGAAATTAAGTCAACCGTTGAGCATCCGATTAACCTGGAAGAGAAATTTCCACAGGTT 1680
QY 1681 CGGCAACTGGCGATGGCCCTGACATATATCTTCTGGGACACGACCGCTTTTGGTGGCTAC 1740
Db 1681 CGGCAACTGGCGATGGCCCTGACATATATCTTCTGGGACACGACCGCTTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATCAACCCGGCAAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGGCTAGCTTACAAACGCAAGCTGATTGCTTACCCGATCCCTGTT 1860
Db 1801 CCGTTTACCTGGGATGGCTAGCTTACAAACGCAAGCTGATTGCTTACCCGATCCCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCAAACCCGCCAAAAACCTTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTATAACAAAGATCTGCTGCCAAACCCGCCAAAAACCTTGGGAA 1920
QY 1921 GAGATCCCGGCTGGATAAAGAACTGAAAGGAAAGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCTGGATAAAGAACTGAAAGCGAAAGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGAACGCTGGGCGTGGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAAAACGCAAGTACGACATTAAGAACGCTGGGCGTGGATTAACGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTGGTGAACCTGATTAATAAACAACAATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTGGTGAACCTGATTAATAAACAACAATGACAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTAATAAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAAGCGGAAACAGCGATGACCATCAAGCGCCGCTGG 2220
QY 2221 GCATGGTCCAAATCGACACACGCAAAAGTGAATTAATGGTGAACCGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACACGCAAAAGTGAATTAATGGTGAACCGTACTGCGGACCTTC 2280
QY 2281 AAGGTCAACCATCAAAACCGTTCGTTGGCGTGTGAGCGCAGGTATTAAACCGCCAGT 2340
Db 2281 AAGGTCAACCATCAAAACCGTTCGTTGGCGTGTGAGCGCAGGTATTAAACCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGACTGATGAAGGTTCTG 2400
QY 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGGGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGGGTAGCGCTGAAGTCTTTACGAGGAAGAG 2460
QY 2461 TTGGCAAAAGATCCACGTTATTCGCCCAACCAATGGAAAAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCAAAAGATCCACGTTATTCGCCCAACCAATGGAAAAACGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTCGCTACTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTCGCTACTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGAGCGCCAGACTTAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACATGTCGATGAAGCCCTGAAAGAGCGCCAGACTTAATTCGAGCTCG 2640
QY 2641 AACAAACAACAATAACAATAACAACCTTCGGGATCGAGGGAAGGATTTCAAGATTC 2700
Db 2641 AACAAACAACAATAACAATAACAACCTTCGGGATCGAGGGAAGGATTTCAAGATTC 2700
QY 2701 GGATCCTCTTCTCTGTGGCCCGCGGCTCTGAGAGCCGG 2741


```
QY 1621 ACCGGAATTAAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAAATTCACACAGGTT 1680
Db 1621 ACCGGAATTAAAGTCAACCGTTGAGCATCCGGATAAACTGGAAGAGAAAATTCACACAGGTT 1680
QY 1681 GCGCAACTGGCGATGGCCCTGACATATCTCTCGGCACACGACCGCTTTGGTGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATATCTCTCGGCACACGACCGCTTTGGTGCTAC 1740
QY 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGGCGTAGCTTACAAACGCAAGCTGATTGCTTACCCGATCCTGTT 1860
Db 1801 CCGTTTACCTGGGATGGCGTAGCTTACAAACGCAAGCTGATTGCTTACCCGATCCTGTT 1860
QY 1861 GAAGGCTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAACCTGGGAA 1920
Db 1861 GAAGGCTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAACCTGGGAA 1920
QY 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGGTAAAGCGCGCTGATGTTCAAC 1980
QY 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATTGCTGCTGACGGGGTTATGCGTTCAAG 2040
QY 2041 TATGAARACGCAAGTACGACATTAAGACGTTGGCGGTGATTAAGCTGGCGGAAAGCG 2100
Db 2041 TATGAARACGCAAGTACGACATTAAGACGTTGGCGGTGATTAAGCTGGCGGAAAGCG 2100
QY 2101 GGTCTGACCTTCTCTGTTGACCTGATTAATAAACAACATCAATGATGACACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTAATAAACAACATCAATGATGACACACCGATTAC 2160
QY 2161 TCCATCGCAGAAAGTCCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCGCTGG 2220
Db 2161 TCCATCGCAGAAAGTCCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCGCTGG 2220
QY 2221 GCATGGTCCACATCGACACGACGAAAGTAAATGATGATGATGATGATGATGATGATGATG 2280
Db 2221 GCATGGTCCACATCGACACGACGAAAGTAAATGATGATGATGATGATGATGATGATGATG 2280
QY 2281 AAGGGTCAACCATCAAAACCGTTCTGTCGCTGAGCGAGGATTAATAACGCCCGCAGT 2340
Db 2281 AAGGGTCAACCATCAAAACCGTTCTGTCGCTGAGCGAGGATTAATAACGCCCGCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGATTCTCTCGAAACTATCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGATTCTCTCGAAACTATCTGACTGATGAAGTCTG 2400
QY 2401 GNAAGCGTTAATAAGACAAACCGCTGGTGGCTAGCGCTAAGTCTTACAGGAAG 2460
Db 2401 GNAAGCGTTAATAAGACAAACCGCTGGTGGCTAGCGCTAAGTCTTACAGGAAG 2460
QY 2461 TTGGCGAAAGATCCACGATTTCGCCCAACCATGGAAGAACCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTTCGCCCAACCATGGAAGAACCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCCGTGCGTACTCGCGTGAATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCCGTGCGTACTCGCGTGAATCAACGCC 2580
QY 2581 GCCAGCGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2640
Db 2581 GCCAGCGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2640
QY 2641 AACAAACAAACAAATTAACAAACCTCGGATCGGATCGGATCGGATCGGATCGGATCGGATCG 2700
Db 2641 AACAAACAAACAAATTAACAAACCTCGGATCGGATCGGATCGGATCGGATCGGATCGGATCG 2700
```

```
QY 2701 GGATCCTCT 2709
Db 2701 GGATCCTCT 2709
```

RESULT 8

```
US-10-343-859-8
; Sequence 8, Application US/10343859
; Publication No. US20040110161A1
; GENERAL INFORMATION:
; APPLICANT: Nanogen Recognition GMBH
; TITLE OF INVENTION: Method for detecting mutations in
; TITLE OF INVENTION: nucleotide sequences
; FILE REFERENCE: 612,406-033
; CURRENT APPLICATION NUMBER: US/10/343,859
; CURRENT FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: PCT/EP01/08127
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 10038237.1
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 6648
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: plasmid
; OTHER INFORMATION: pMAL-c2x
US-10-343-859-8
```

```
Query Match 81.8%; Score 2701; DB 19; Length 6648;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2704; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGTATCGCATGATAGCCCGGAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGTATCGCATGATAGCCCGGAGAGA 60
QY 61 GTCATTCAGGGTGTGATGTGAAACCCAGTAAAGCTTATAGCATGTCGACAGATGCGG 120
Db 61 GTCATTCAGGGTGTGATGTGAAACCCAGTAAAGCTTATAGCATGTCGACAGATGCGG 120
QY 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACACGCGCAGCCAGCTTCTCGAAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACACGCGCAGCCAGCTTCTCGAAAA 180
QY 181 CGCGGAAAGTGGNAGCGGGGATGGCGAGCTGAATTAATCCCAACCGCGTGGCAC 240
Db 181 CGCGGAAAGTGGNAGCGGGGATGGCGAGCTGAATTAATCCCAACCGCGTGGCAC 240
QY 241 AACAACTCGCGGCAAAACAGTCTGTTGTTGGGTTGGCCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTCGCGGCAAAACAGTCTGTTGTTGGGTTGGCCACCTCCAGTCTGGCCCTGC 300
QY 301 ACGGCGCTCGCAAAATGTCGCGCGATTAATATCTCGCGCGATCAATCTGGGTGCGAG 360
Db 301 ACGGCGCTCGCAAAATGTCGCGCGATTAATATCTCGCGCGATCAATCTGGGTGCGAG 360
QY 361 TGGTGTCTCGATGCTAGAACGAGCGGCTGAGAGCTGTAAAGGGGGTGCACATC 420
Db 361 TGGTGTCTCGATGCTAGAACGAGCGGCTGAGAGCTGTAAAGGGGGTGCACATC 420
QY 421 TTCTCGCCAAACGCGTCAAGTGGCTGATCAATTAATCTATCCGCTGATGACCAAGTGC 480
Db 421 TTCTCGCCAAACGCGTCAAGTGGCTGATCAATTAATCTATCCGCTGATGACCAAGTGC 480
QY 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTATTTCTTGATGTTCTGACCA 540
Db 481 TTGCTGTGGAAGCTGCTGCACTAATGTTCCGGCGTATTTCTTGATGTTCTGACCA 540
QY 541 CACCCATCAACGATTAATTTTCTCCATGAAGAGCGTACCGGACTGGCGGTGGAGCATC 600
Db 541 CACCCATCAACGATTAATTTTCTCCATGAAGAGCGTACCGGACTGGCGGTGGAGCATC 600
```

Db 541 CACCACATCAACAGTATTATTTTCTCCCATGAAGACGGTACGCGACTGGCGGTGGAGCATC 600
Qy TGGTCGCAATTGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGCTCGG 660
Db TGGTCGCAATTGGGTACACAGCAAAATCGCGCTGTAGCGGGCCCATTAAGTTCTGCTCGG 660
Qy GCGGCTCGGCTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 720
Db GCGGCTCGGCTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 720
Qy CGGAACGGGAAGGACATCGGATGCCATGTCCGTTTCAACAAACCAATGCAATGCTGA 780
Db CGGAACGGGAAGGACATCGGATGCCATGTCCGTTTCAACAAACCAATGCAATGCTGA 780
Qy ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTCGCAACGATCAGATGCGCTGGGCGCAA 840
Db ATGAGGGCATCGTTCCTCACTGCGATGCTGGTTCGCAACGATCAGATGCGCTGGGCGCAA 840
Qy TGGCGGCCATTACCGAGTCCGGGCTGCGGCTGGTGGTGGGATATCTCGGTAGTGGGATAG 900
Db TGGCGGCCATTACCGAGTCCGGGCTGCGGCTGGTGGTGGGATATCTCGGTAGTGGGATAG 900
Qy ACGATACCGGAACAGAGTCATGTTATATCCGCGGTTAACCCACCATCAAAACAGATTTTC 960
Db ACGATACCGGAACAGAGTCATGTTATATCCGCGGTTAACCCACCATCAAAACAGATTTTC 960
Qy GCCTGCTGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db GCCTGCTGGGCAACACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Qy AGGCAATCAGGTGTTGCCGTCTCATCTGGTGAAGAAACCAACCCCTGGCGGCCAATA 1080
Db AGGCAATCAGGTGTTGCCGTCTCATCTGGTGAAGAAACCAACCCCTGGCGGCCAATA 1080
Qy CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATCGAGTGGCAGCAGGTTT 1140
Db CGCAACCGCTCTCCCGCGGTTGGCGGATTCATTAATCGAGTGGCAGCAGGTTT 1140
Qy CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Db CCCGACTGGAAGCGGCGAGTGAGCGCAACGCAATTAATGTGAGTTAGCTCACTCATTTAG 1200
Qy GCACAAATCTCATGTTTGACAGCTTATCATCGATGCAACGCTGCAACCAATGCTTCTGGCG 1260
Db GCACAAATCTCATGTTTGACAGCTTATCATCGATGCAACGCTGCAACCAATGCTTCTGGCG 1260
Qy TCAGGAGCCATCGAAGCTGTGGTATGCGTGTGCGTGTGAGTCACTGCATAATTCG 1320
Db TCAGGAGCCATCGAAGCTGTGGTATGCGTGTGCGTGTGAGTCACTGCATAATTCG 1320
Qy TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCATAAACGGTT 1380
Db TGTGCTCAAGCGGCACTCCCGTTCTGGATAATGTTTTTTCGCGCGACATCATAAACGGTT 1380
Qy CTGGCAATATCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATTAATGTGGGA 1440
Db CTGGCAATATCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTATTAATGTGGGA 1440
Qy ATTGTGAGCGGATAACAAATTTACACAGGAACAGCAGTCCGTTTATAGGTGTTTTTCAGGA 1500
Db ATTGTGAGCGGATAACAAATTTACACAGGAACAGCAGTCCGTTTATAGGTGTTTTTCAGGA 1500
Qy GCATTTCAACCAAGGACCATAGATATTGAAACTGAAGAGGTAAACTGGTAATCTGG 1560
Db GCATTTCAACCAAGGACCATAGATATTGAAACTGAAGAGGTAAACTGGTAATCTGG 1560
Qy ATTAACGGCGATAAGGCTATAACGGTCTCGCTGAAGTCCGTAAGAAATTCGAGAAAGAT 1620
Db ATTAACGGCGATAAGGCTATAACGGTCTCGCTGAAGTCCGTAAGAAATTCGAGAAAGAT 1620
Qy ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCACAGGTT 1680
Db ACCGGAATTAAGTCAACGGTTGAGCATCCGGATAAACTGGAAGAGAAATTCGCCACAGGTT 1680

Qy 1681 GCGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCACACGACCCGTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTTATCTTCTGGGCACACGACCCGTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCCTGTTGGCTGAAATCACCCCGGACAAAGCGTTCCAGGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTTAAACGGAAGCTGATTTGCTTACCCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTAAACGGAAGCTGATTTGCTTACCCGATCGCTGTT 1860
Qy 1861 GAAGCGTTATCGCTGATTTATACAAAGATCTGCTGCGAACCCGCCCAAAACCTGGGAA 1920
Db 1861 GAAGCGTTATCGCTGATTTATACAAAGATCTGCTGCGAACCCGCCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCGCTGGATAAAGAACTCAAAAGCGAAAGTAAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAAGAACTCAAAAGCGAAAGTAAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTTGCTGCTGACCGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTTCACCTGCGCGCTGATTTGCTGCTGACCGGGGTTATGCGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTTAAAGACGTGGCGGTGATTAACGCTGGCGGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTTAAAGACGTGGCGGTGATTAACGCTGGCGGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAAACAACACATGAATGCAGACACCGATTC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTTAAAAAACAACACATGAATGCAGACACCGATTC 2160
Qy 2161 TCCATCGCAAGAGTGGCTTTTAAATAAAGGCAACACGCGATGACCATCAACCGGCCGTGG 2220
Db 2161 TCCATCGCAAGAGTGGCTTTTAAATAAAGGCAACACGCGATGACCATCAACCGGCCGTGG 2220
Qy 2221 GCATGCTCAACATCGACACGCAAGTCAATTTATGCTGTAAACGCTGCTGCGACCTTC 2280
Db 2221 GCATGCTCAACATCGACACGCAAGTCAATTTATGCTGTAAACGCTGCTGCGACCTTC 2280
Qy 2281 AAGGTCACCAATCCAAACCGTTTCTGAGCGCAGGTATTTAACGCGCCAGT 2340
Db 2281 AAGGTCACCAATCCAAACCGTTTCTGAGCGCAGGTATTTAACGCGCCAGT 2340
Qy 2341 CCGAACAAAGAGTGGCAAAAGAGTTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
Db 2341 CCGAACAAAGAGTGGCAAAAGAGTTTCTCGAAAACTATCTGCTGACTGATGAAGGTCG 2400
Qy 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Qy 2461 TTGGCGAAAGATCCAGTATTTGCCGCAACATGGAAGAACCCAGAAAGGTGAAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTTGCCGCAACATGGAAGAACCCAGAAAGGTGAAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGGGTGATCAAGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCGTACTGGGTGATCAAGCC 2580
Qy 2581 GCCAGCGGTGCTCAGACTGTTCGATGAAGCCCTGAAAGACCGCGCAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTGCTCAGACTGTTCGATGAAGCCCTGAAAGACCGCGCAGACTAATTCGAGCTCG 2640
Qy 2641 AACACAAACAAATTAACAAATTAACAAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACACAAACAAATTAACAAATTAACAAACCTCGGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Qy 2701 GGATCTCT 2709
Db 2701 GGATCTCT 2709

RESULT 9

US-10-343-859-9
; Sequence 9, Application US/10343859
; Publication No. US20040110161A1
; GENERAL INFORMATION:
; APPLICANT: Nanogen Recognomics GMBH
; TITLE OF INVENTION: Method for detecting mutations in
; FILE OF INVENTION: nucleotide sequences
; FILE REFERENCE: 612,406-033
; CURRENT APPLICATION NUMBER: US/10/343,859
; PRIOR FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: PCT/EP01/08127
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 10038237.1
; PRIOR FILING DATE: 2000-08-04
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 9191
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Description of the artificial sequence: pMALc2x-muts
; OTHER INFORMATION: Plasmid
US-10-343-859-9

Query Match 81.8%; Score 2698; DB 19; Length 9191;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2701; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGGCCCGGGAAGAGA 60
DB 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGGCCCGGGAAGAGA 60

QY 61 GTCAATTCAGGTGGTGAATGTGAACACAGTAACGTTATACCATGTCGACAGATGCGG 120
DB 61 GTCAATTCAGGTGGTGAATGTGAACACAGTAACGTTATACCATGTCGACAGATGCGG 120

QY 121 GTGTCTCTTATCAGACCGGTTTCCCGCGTGGTGAACACAGGCCAGCCACGTTTCTGGGAAAA 180
DB 121 GTGTCTCTTATCAGACCGGTTTCCCGCGTGGTGAACACAGGCCAGCCACGTTTCTGGGAAAA 180

QY 181 CCGGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTAATTCACACCGCGTGGCAC 240
DB 181 CCGGGAAAAAGTGAAGCGGCGATGGCGAGCTGAATTAATTCACACCGCGTGGCAC 240

QY 241 AACAACTGGCGGGCAACAGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300
DB 241 AACAACTGGCGGGCAACAGTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC 300

QY 301 ACGCGCGTCCGCAATTTGTCCGGCGATTAAATCTCGCGCCGATCACTGGGTGGCAGCG 360
DB 301 ACGCGCGTCCGCAATTTGTCCGGCGATTAAATCTCGCGCCGATCACTGGGTGGCAGCG 360

QY 361 TGGTGTGTGATGTAGAACGAAGCGGCTCGAAGCCTGTAAAGCGGCGGTGCACAATC 420
DB 361 TGGTGTGTGATGTAGAACGAAGCGGCTCGAAGCCTGTAAAGCGGCGGTGCACAATC 420

QY 421 TTCTCCGCAACCGCTGATGGGTGATCAATTAATCTATCCGTGGATGACAGGATGCCA 480
DB 421 TTCTCCGCAACCGCTGATGGGTGATCAATTAATCTATCCGTGGATGACAGGATGCCA 480

QY 481 TTGCTGTGGAAGCTGCTCAGCTAATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA 540
DB 481 TTGCTGTGGAAGCTGCTCAGCTAATGTTCCGGCGTTATTTCTGATGCTCTGACCCAGA 540

QY 541 CACCCATCAACAGTATATTTTCTCCCATGAAGACCGTACCGACTGGCGGTGGAGCATC 600
DB 541 CACCCATCAACAGTATATTTTCTCCCATGAAGACCGTACCGACTGGCGGTGGAGCATC 600

QY 601 TGGTCCATTTGGGTCAACAGCAAAATCCGCTGTATAGCGGCCCATTAAGTTCTGTCTCGG 660
DB 601 TGGTCCATTTGGGTCAACAGCAAAATCCGCTGTATAGCGGCCCATTAAGTTCTGTCTCGG 660

DB 601 TGGTCCATTTGGGTCAACAGCAAAATCCGCTGTATAGCGGCCCATTAAGTTCTGTCTCGG 660
QY 661 CGCTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
DB 661 CGCTCTGCGTCTGCGTGGCATAAATATCTCACTCGCAATCAAAATTCAGCGGATAG 720
QY 721 CGGAACGGGAAGCGCACTGGAGTGCCTATGTCGGGTTTCAACAAACCAATGCAAAATGCTGA 780
DB 721 CGGAACGGGAAGCGCACTGGAGTGCCTATGTCGGGTTTCAACAAACCAATGCAAAATGCTGA 780
QY 781 ATGAGGCGCATCGTTCCCACTGCGATGCTGTTGCCAAGATCAGATGCGCTGGGCGCAA 840
DB 781 ATGAGGCGCATCGTTCCCACTGCGATGCTGTTGCCAAGATCAGATGCGCTGGGCGCAA 840
QY 841 TGGCGGCCATTTACGAGTCCGGGCTGCGGTTGGTGGGATATCTCGTAGTGGGATACG 900
DB 841 TGGCGGCCATTTACGAGTCCGGGCTGCGGTTGGTGGGATATCTCGTAGTGGGATACG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAACAGGATTTTC 960
QY 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGTGA 1020
DB 961 GCCTGCTGGGGCAAAACAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGTGA 1020
QY 1021 AGGCAATCAGCTGTGCTCTCATCTGCTGTAAGAAAGAAACCAACCTCGGCGCCAAATA 1080
DB 1021 AGGCAATCAGCTGTGCTCTCATCTGCTGTAAGAAAGAAACCAACCTCGGCGCCAAATA 1080
QY 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGATTCATTAATGCAGTGGCAGCAGCAGGTTT 1140
DB 1081 CGCAAAACCGCTCTCCCGCGGTTGGCGATTCATTAATGCAGTGGCAGCAGCAGGTTT 1140
QY 1141 CCCGACTGGAAAGGGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
DB 1141 CCCGACTGGAAAGGGGCGAGTGAGCGCAACGCAATTAATGTAGTTAGTCACTCATTTAG 1200
QY 1201 GCACAATTTCTCATGTTGACAGCTTATCATCGACTGCAAGCGTGCACCAATGTTCTGGCG 1260
DB 1201 GCACAATTTCTCATGTTGACAGCTTATCATCGACTGCAAGCGTGCACCAATGTTCTGGCG 1260
QY 1261 TCAGCAGCCATCGGAAGCTGTGTTATGCTGTCGCGTCTGAATCACTCATTAATTCG 1320
DB 1261 TCAGCAGCCATCGGAAGCTGTGTTATGCTGTCGAGTCTGAATCACTCATTAATTCG 1320
QY 1321 TGTGCTCAAGCGCAGCTCCCGTCTTGATTAATGTTTTTGGCGCGACATCATACCGTT 1380
DB 1321 TGTGCTCAAGCGCAGCTCCCGTCTTGATTAATGTTTTTGGCGCGACATCATACCGTT 1380
QY 1381 CTGGCAAAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
DB 1381 CTGGCAAAATATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA 1440
QY 1441 ATTGTAGCGGATTAACAATTTTCAACAGGAACAGCCAGTCCGTTAGTGTTCACGA 1500
DB 1441 ATTGTAGCGGATTAACAATTTTCAACAGGAACAGCCAGTCCGTTAGTGTTCACGA 1500
QY 1501 GCACCTTCCACCAACAGGACCATAGATTATGAACCTGAAGAGGTAATACTGTAATCTGG 1560
DB 1501 GCACCTTCCACCAACAGGACCATAGATTATGAACCTGAAGAGGTAATACTGTAATCTGG 1560
QY 1561 ATTAACGGCGATTAAGGCTATAACCGTCTCGCTGAAGTCCGTAAGAAATTCAGAAAGAT 1620
DB 1561 ATTAACGGCGATTAAGGCTATAACCGTCTCGCTGAAGTCCGTAAGAAATTCAGAAAGAT 1620
QY 1621 ACCGAAATTAAGTCAACCGTTGAGCATCCGGATAAATGGAAGAGAAATTCACAGGTT 1680
DB 1621 ACCGAAATTAAGTCAACCGTTGAGCATCCGGATAAATGGAAGAGAAATTCACAGGTT 1680
QY 1681 GCGGCAATCGGCGATGGCCCTGACATTAATCTCTGGGCACACGCGCTTTCTGCTGCTAC 1740
DB 1681 GCGGCAATCGGCGATGGCCCTGACATTAATCTCTGGGCACACGCGCTTTCTGCTGCTAC 1740

```
QY 1741 GCTCAATCTGGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGCAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAAATCACCCCGGACAAAGCGTTCCAGCAAGCTGTAT 1800
QY 1801 CCGTTTACTGGATGCCGTAGCTTTACAAACGCAAGCTGATGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACTGGATGCCGTAGCTTTACAAACGCAAGCTGATGCTTACCCGATCGCTGT 1860
QY 1861 GAAGGTTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCCGCGCAAAACCTGGAA 1920
Db 1861 GAAGGTTTATCGCTGATTTATAACAAGATCTGCTGCCGAACCCGCGCAAAACCTGGAA 1920
QY 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAGCGAAAGTAAAGAGCGCGCTGATTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATTAAGAACTGAAGCGAAAGTAAAGAGCGCGCTGATTTCAAC 1980
QY 1981 CTGAAGAACCTGACTTACCTGCGCGCTGATGCTGCTGAGCGGGGTATGCTTCAAG 2040
Db 1981 CTGAAGAACCTGACTTACCTGCGCGCTGATGCTGCTGAGCGGGGTATGCTTCAAG 2040
QY 2041 TATGAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATACGCTGGCGCAAAAGCG 2100
Db 2041 TATGAAAACGCAAGTACGACATTAAGACGCTGGCGTGGATACGCTGGCGCAAAAGCG 2100
QY 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACAACATGAATGAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCGTTGACCTGATTAATAAACAACAACATGAATGAGACACCGATTAC 2160
QY 2161 TCCATCCGAGAGCTGCTTTAATAAGGCGAAACAGCGATGACATCAACGGCCGCTGG 2220
Db 2161 TCCATCCGAGAGCTGCTTTAATAAGGCGAAACAGCGATGACATCAACGGCCGCTGG 2220
QY 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTAATGCTGTAACCGTACTGCCACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTAATGCTGTAACCGTACTGCCACCTTC 2280
QY 2281 AAGGGTCAACCATCAAAACCGTTGTTGGCTGCTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGGTCAACCATCAAAACCGTTGTTGGCTGCTGAGCGCAGGTATTAACGCCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACCTATCTGCTGACTGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAGACAAACCGCTGGCTGCGCTAGCGCTGAAGTCTTACGAGGAAGAG 2460
Db 2401 GAAGCGGTTAATAAGACAAACCGCTGGCTGCGCTAGCGCTGAAGTCTTACGAGGAAGAG 2460
QY 2461 TTGGCGAAAGATCCAGCTATTGCGCGCACCAATGAAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGCTATTGCGCGCACCAATGAAACCGCCAGAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCGCGAGATGTCGGCTTTCTGGTATGCGGTGCTGCTGCGGTGATCAACGCC 2580
Db 2521 CCGAACATCCGCGAGATGTCGGCTTTCTGGTATGCGGTGCTGCTGCGGTGATCAACGCC 2580
QY 2581 GCCAGCGGTCTGACAGCTGCTGAAGAGCCCTGAAAGACGCGAGACTAATTCGAGCTCG 2640
Db 2581 GCCAGCGGTCTGACAGCTGCTGAAGAGCCCTGAAAGACGCGAGACTAATTCGAGCTCG 2640
QY 2641 AACCAACAAACAATAACAATAACAACAACTCGGGATCGAGGGAAGATTTTCAGAATTC 2700
Db 2641 AACCAACAAACAATAACAATAACAACAACTCGGGATCGAGGGAAGATTTTCAGAATTC 2700
QY 2701 GGATCC 2706
Db 2701 GGATCC 2706
```

RESULT 10

US-10-263-153-40

J Sequence 40, Application US/10263153

```
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984 US.01
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40
; LENGTH: 7112
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3189)
; OTHER INFORMATION: pMBP-c2X-Toxop30del11 (52-214aa)
; US-10-263-153-40
```

```
Query Match 81.5%; Score 2690.4; DB 18; Length 7112;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
```

```
QY 1 CCGACACCATCGAATGTTGCAAAACCTTTCCGGTATGGCATGATAGCGCCCGAAGAGA 60
Db 1 CCGACACCATCGAATGTTGCAAAACCTTTCCGGTATGGCATGATAGCGCCCGAAGAGA 60
QY 61 GTCAATTCAGGGTGGTGAATGTGAAACCAAGTATACGATGTACGATGTCCGAGTATGCC 120
Db 61 GTCAATTCAGGGTGGTGAATGTGAAACCAAGTATACGATGTACGATGTCCGAGTATGCC 120
QY 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACCAAGCCAGCCAGCGTTCTTCGCAAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCCGGTGGTGAACCAAGCCAGCCAGCGTTCTTCGCAAAA 180
QY 181 CCGGGGAAAAAGTGGAAAGCGCGCATGGCGAGCTGAATTACATTCCCAACCGCGTGGCAC 240
Db 181 CCGGGGAAAAAGTGGAAAGCGCGCATGGCGAGCTGAATTACATTCCCAACCGCGTGGCAC 240
QY 241 AACAACTGGCGGCAACACAGTCGTGCTGATTTCGCGTTCGCCACTCCAGTCTGCCCTGC 300
Db 241 AACAACTGGCGGCAACACAGTCGTGCTGATTGGCGTTGCCACTCCAGTCTGCCCTGC 300
QY 301 ACGCCCGCTCGCAAAATTTGTCGCGCGATTAATAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCCCGCTCGCAAAATTTGTCGCGCGATTAATAATCTCGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGTGTCGATGCTAGAACGAAGCGCGTCAAGCTGTAAAGCGCGGTGGCAATC 420
Db 361 TGGTGTGTCGATGCTAGAACGAAGCGCGTCAAGCTGTAAAGCGCGGTGGCAATC 420
QY 421 TTCTCGCGCAACCGCTCAGTGGCTCATCATTAATCATTCGCTGGATGACCAAGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGCTCATCATTAATCATTCGCTGGATGACCAAGATGCCA 480
QY 481 TTGCTGTGAAGCTGCTGCACATAATGTTCCGCGTTATTTCTTGTATGCTCTGACCCAGA 540
Db 481 TTGCTGTGAAGCTGCTGCACATAATGTTCCGCGTTATTTCTTGTATGCTCTGACCCAGA 540
QY 541 CACCCATCAACAGTATTTTCTCCATGAAGACCGTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTTCTCCATGAAGACCGTACGCGACTGGGCGTGGAGCATC 600
QY 601 TGGTCCGATTTGGGTACACAGCAATCGCTGTTAGCGGCGCCATTAAGTTCTGCTCGG 660
Db 601 TGGTCCGATTTGGGTACACAGCAATCGCTGTTAGCGGCGCCATTAAGTTCTGCTCGG 660
QY 661 CCGGTCTGCGTCTGGCTGGCATAAATAATCTCACTCGCAATCAAAATTCAGCCGATAG 720
```

Db 661 |||||CGGCTCTCGGCTGGCTGGCAATAATCTCCTCGCAATCAAAATTCAGCCGATAG 720
Qy 721 CGGAACGGGAAGGACATGGAGTGCATGTCCGGTTTTCAACAAACCAATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGGACATGGAGTGCATGTCCGGTTTTCAACAAACCAATGCAAAATGCTGA 780
Qy 781 ATGAGGGATCTGTTCCCATCTGGATGCTGGTTGCCAACAGATCAGATGCGCTGGCGCAA 840
Db 781 ATGAGGGATCTGTTCCCATCTGGATGCTGGTTGCCAACAGATCAGATGCGCTGGCGCAA 840
Qy 841 TGGCGCCATTACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGTAGTGGGATACG 900
Db 841 TGGCGCCATTACCGAGTCCGGCTGCGGTTGGTGGGATATCTCGTAGTGGGATACG 900
Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCAATCAAAACAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCAATCAAAACAGGATTTTC 960
Qy 961 GCCTGCTGGGCAACACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGTGA 1020
Db 961 GCCTGCTGGGCAACACGAGCTGGACCGCTTGCTGCAACTCTCTCAGGGCCAGGCGTGA 1020
Qy 1021 AGGGCAATCAGCTGTTGCCGCTCTCACTGCTGAAAGAAACCAACCTGCGCCCAATA 1080
Db 1021 AGGGCAATCAGCTGTTGCCGCTCTCACTGCTGAAAGAAACCAACCTGCGCCCAATA 1080
Qy 1081 CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTAAATGAGCTGGCAGCACAGGTTT 1140
Db 1081 CGCAACCGCCTCTCCCGCGGTTGGCGGATTCATTAAATGAGCTGGCAGCACAGGTTT 1140
Qy 1141 CCCGACTGGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTAGTCACTCATTTAG 1200
Db 1141 CCCGACTGGGAAGCGGCGAGTGAGCGCAACGCAATTAATGAGTAGTCACTCATTTAG 1200
Qy 1201 GCACATTTCTATGTTTCACAGCTTATCATGACCTGCAGCTGCAGCAATCTCTCTGGG 1260
Db 1201 GCACATTTCTATGTTTCACAGCTTATCATGACCTGCAGCTGCAGCAATCTCTCTGGG 1260
Qy 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGCAGCTCGTAATCACTGCATAAATTCG 1320
Db 1261 TCAGGACGCCATCGGAAGCTGTGGTATGGCTGTGCAGCTCGTAATCACTGCATAAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACATCCCGTTCTGATATGTTTTTTCGCGCACATCAATACGTT 1380
Db 1321 TGTGCTCAAGCGCACATCCCGTTCTGATATGTTTTTTCGCGCACATCAATACGTT 1380
Qy 1381 CTGGCAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
Db 1381 CTGGCAATATCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
Qy 1441 ATTGTAGCGGATACAAATTTTACACAGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Db 1441 ATTGTAGCGGATACAAATTTTACACAGAAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Qy 1501 GCATTTCAACCAAGGACCATAGCATATGAAATCGAAGAGGTAAATCGTAAATCTGG 1560
Db 1501 GCATTTCAACCAAGGACCATAGCATATGAAATCGAAGAGGTAAATCGTAAATCTGG 1560
Qy 1561 ATTAACGGCGATAAGGCTATACCGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAGGCTATACCGTCTCGCTGAAGTCGGTGAAGAAATTCGAGAAAGAT 1620
Qy 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTGAGCATCCGGATAAATCGGAAGAGAAATTCACAGGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTCTGGGCACACGACCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGCTGCTGAAATCAACCGGCAAAAGCGTTCCAGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGCTGCTGAAATCAACCGGCAAAAGCGTTCCAGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTGCTTACCCGATCGCTGT 1860
Db 1801 CCGTTTACCTGGGATGCGGTACGTTTACAAACGCAAGCTGATTGCTTACCCGATCGCTGT 1860
Qy 1861 GAACGGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAACCTGGGAA 1920
Db 1861 GAACGGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAACCCGCCAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCGCTGGAATAAGAACTGAAAGCGAAAGTGAAGAGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGAATAAGAACTGAAAGCGAAAGTGAAGAGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGTAACGGGGTTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGCGCGCTGATGCTGTAACGGGGTTATGCGTTCAAG 2040
Qy 2041 TATCAAAACGCAAGTACGACATTTAAAGACGTGGCGTGGATAACGCTGGCGGCAAGCG 2100
Db 2041 TATCAAAACGCAAGTACGACATTTAAAGACGTGGCGTGGATAACGCTGGCGGCAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGGTTGACCTGATTAATAAACAACACATGAATGCAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGGCCGCTGG 2220
Db 2161 TCCATCGCAGAAGCTGCTTTAATAAGCGGAAACAGCGATGACCATCAACGGGCCGCTGG 2220
Qy 2221 GCATGCTTCAACATCGACACCGCAAGTGAATTAATGCTGTAACGGTAAACGGACCTTC 2280
Db 2221 GCATGCTTCAACATCGACACCGCAAGTGAATTAATGCTGTAACGGTAAACGGACCTTC 2280
Qy 2281 AAGGTCACACCATCAACCGTTCGTTGGCGTGTGAGCGAGGTATTAACCGGCCAGT 2340
Db 2281 AAGGTCACACCATCAACCGTTCGTTGGCGTGTGAGCGAGGTATTAACCGGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACCTATCTGCTGACTGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAACCTATCTGCTGACTGATGAAGTCTG 2400
Qy 2401 GAACGGTTAATAAGACAAACCGCTGGTGGCTAGCGCTGAAAGTCTTACGAGGAAG 2460
Db 2401 GAACGGTTAATAAGACAAACCGCTGGTGGCTAGCGCTGAAAGTCTTACGAGGAAG 2460
Qy 2461 TTGGCGAAAGATCCACGTTATTCGCCCACTATGGAACCGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGTTATTCGCCCACTATGGAACCGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTATGCGCTGCTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGTTATGCGCTGCTGCTGCTGATCAACGCC 2580
Qy 2581 GCGAGCGTCTGACAGCTGCTGATGAAGCCCTGAAAGACGCGCAGACATTAATTCGAGTCTG 2640
Db 2581 GCGAGCGTCTGACAGCTGCTGATGAAGCCCTGAAAGACGCGCAGACATTAATTCGAGTCTG 2640
Qy 2641 AACCAACCAACAAATAACAATAACAACCTCGGATCGAGGGAAGGATTTTCAGAAATTC 2700
Db 2641 AACCAACCAACAAATAACAATAACAACCTCGGATCGAGGGAAGGATTTTCAGAAATTC 2700

RESULT 11
US-10-263-153-30
; Sequence 30, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.

; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; FILE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 7259
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3336)
; OTHER INFORMATION: pMBP-c2X-Toxop30del14del18 (83-294aa)
US-10-263-153-30

Query Match 81.5%; Score 2690.4; DB 18; Length 7259;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAGAAACCTTTCCGGGTATGGCATGATAGCGCCCGAAGAGA 60

Qy 61 GTCAAATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTGCGAGAGTATGCCG 120
Db 61 GTCAAATTCAGGGTGGTGAATGTGAACACAGTAACGTTATACGATGTGCGAGAGTATGCCG 120

Qy 121 GTGCTCTTATCAGACCGTTTCCCGCGTGGTGAACACCGAGTGAATTTCCGAAAAA 180
Db 121 GTGCTCTTATCAGACCGTTTCCCGCGTGGTGAACACCGAGTGAATTTCCGAAAAA 180

Qy 181 CGCGGAAAAAGTGGAAAGCGCGGATGGCGAGCTGAATTAATTTCCAAACCGCGTGGCAC 240
Db 181 CGCGGAAAAAGTGGAAAGCGCGGATGGCGAGCTGAATTAATTTCCAAACCGCGTGGCAC 240

Qy 241 AACAACTGGCGGGGAAACAGTCGTTGTGATGTGGCTGGTGCACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGCGGGGAAACAGTCGTTGTGATGTGGCTGGTGCACCTCCAGTCTGGCCCTGC 300

Qy 301 ACGCGCGTGCAGAAATGTTCGCGCGATTAATCTCGCGCGATCAACTGGTGGCCAGCG 360
Db 301 ACGCGCGTGCAGAAATGTTCGCGCGATTAATCTCGCGCGATCAACTGGTGGCCAGCG 360

Qy 361 TGGTGGTTCGATGGTGAAGAGCGCGTGAAGCTCTGAAGCGTGAAGCGCGGTGCACATC 420
Db 361 TGGTGGTTCGATGGTGAAGAGCGCGTGAAGCTCTGAAGCGCGGTGCACATC 420

Qy 421 TTCTCGCGCAACCGCTCAGTGGGTGATCATTAATCTATCCGCTGGATGACAGGATGCCA 480
Db 421 TTCTCGCGCAACCGCTCAGTGGGTGATCATTAATCTATCCGCTGGATGACAGGATGCCA 480

Qy 481 TTGCTGTGGAAGCTCGCTGACATTAATCTCGCGGTATTTCTTGATGTCTTGACACAGA 540
Db 481 TTGCTGTGGAAGCTCGCTGACATTAATCTCGCGGTATTTCTTGATGTCTTGACACAGA 540

Qy 541 CACCCATCAACAGTATTTTCTCCCATGAAGACGCTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTTTCTCCCATGAAGACGCTACGCGACTGGGCGTGGAGCATC 600

Qy 601 TGGTGGCATTTGGGTACACAGCAATCGCGTGTAGCGGGCCCATTAAGTTCGTCTCGG 660
Db 601 TGGTGGCATTTGGGTACACAGCAATCGCGTGTAGCGGGCCCATTAAGTTCGTCTCGG 660

Qy 661 CGGCTCTGCGCTGCGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCCATAG 720
Db 661 CGGCTCTGCGCTGCGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCCATAG 720

Qy 721 CGGAAACGGGAAGCGGACTGAGTGCATGTCGGTTCGAAACCAACCATGCAATGCTGA 780
Db 721 CGGAAACGGGAAGCGGACTGAGTGCATGTCGGTTCGAAACCAACCATGCAATGCTGA 780

Qy 781 ATGAGGCGATCGTTCCCACTGCGATGCTGGTCCAAACGATCAGATCGCGCTGGCGCAA 840
Db 781 ATGAGGCGATCGTTCCCACTGCGATGCTGGTGGCAAACGATCAGATCGCGCTGGCGCAA 840

Qy 841 TGGCGCCATTACCGAGTCCGGGCTCGCGTGGTGGCGATATCTCGGTAGTGGGATACG 900
Db 841 TGGCGCCATTACCGAGTCCGGGCTCGCGTGGTGGCGATATCTCGGTAGTGGGATACG 900

Qy 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAACCAACATCAACAGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCGCGGTTAACCAACATCAACAGATTTTC 960

Qy 961 GCCTGCTGGGCGAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020
Db 961 GCCTGCTGGGCGAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGGCCAGCGGTGA 1020

Qy 1021 AGGCAATACGTGTTGCCGCTCTCACTGGTGAAGAAAAAACCCCTCGCGCCCAATA 1080
Db 1021 AGGCAATACGTGTTGCCGCTCTCACTGGTGAAGAAAAAACCCCTCGCGCCCAATA 1080

Qy 1081 CGCAACCGCTCTCCCGCGCGTGGCCGATTCAATTAATGACGCTGGGACGACAGTTT 1140
Db 1081 CGCAACCGCTCTCCCGCGCGTGGCCGATTCAATTAATGACGCTGGGACGACAGTTT 1140

Qy 1141 CCGGACTGGAAAGCGGCGAGTGAAGCAACGCAATTAATGAGTTAGCTCACTATTAG 1200
Db 1141 CCGGACTGGAAAGCGGCGAGTGAAGCAACGCAATTAATGAGTTAGCTCACTATTAG 1200

Qy 1201 GCAAAATTCATGTTTGAACAGCTTATCATCGACTGCACGGTGCACCAATGCTTGGCG 1260
Db 1201 GCAAAATTCATGTTTGAACAGCTTATCATCGACTGCACGGTGCACCAATGCTTGGCG 1260

Qy 1261 TCAGGCGACCATCGGAAGCTGTGATGGCTGTGCGAGTGTAAATCAGTCGCAATTCG 1320
Db 1261 TCAGGCGACCATCGGAAGCTGTGATGGCTGTGCGAGTGTAAATCAGTCGCAATTCG 1320

Qy 1321 TGTGCTCAAGCGCGACTCCCGCTTCTGGATAATGTTTTTTCGCCGACATCAACCGTT 1380
Db 1321 TGTGCTCAAGCGCGACTCCCGCTTCTGGATAATGTTTTTTCGCCGACATCAACCGTT 1380

Qy 1381 CTGCGCAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440
Db 1381 CTGCGCAATTTCTGAAATGAGCTGTGACAAATTAATCATCGGCTCGTAAATGTTGGA 1440

Qy 1441 ATTGTGAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTTTTCACGA 1500
Db 1441 ATTGTGAGCGGATAACAAATTTTACACAGGAAACAGCCAGTCCGTTTAGGTTTTCACGA 1500

Qy 1501 GCACCTTCAACCAACAGGACCATAGATTATGAAACTGAAGAAAGGTAAACTGGTAATCTGG 1560
Db 1501 GCACCTTCAACCAACAGGACCATAGATTATGAAACTGAAGAAAGGTAAACTGGTAATCTGG 1560

Qy 1561 ATTAAACGGGATTAAGGCTTAACCGTCTCGTGAAGTGGTAAAGAAATTCGAGAAAGAT 1620
Db 1561 ATTAAACGGGATTAAGGCTTAACCGTCTCGTGAAGTGGTAAAGAAATTCGAGAAAGAT 1620

Qy 1621 ACCGGAATTAAGTCAACCGTTCGAGTATCCGATTAACCTGGAAGAAATTCGAGAGTT 1680
Db 1621 ACCGGAATTAAGTCAACCGTTCGAGTATCCGATTAACCTGGAAGAAATTCGAGAGTT 1680

Qy 1681 GGGCAACTGGGATGGCCCTGACATTATCTTCTGGGCAACGACCGCTTGGTGGCTAC 1740
Db 1681 GGGCAACTGGGATGGCCCTGACATTATCTTCTGGGCAACGACCGCTTGGTGGCTAC 1740

Qy 1741 GCTCAATCTGGCTGTGGCTGAATCAACCGGCAAAAGCGTTCAGGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTGGCTGAATCAACCGGCAAAAGCGTTCAGGACAAAGCTGTAT 1800

Qy 1801 CGGTTTACCTGGGATGCGCTAGCTTACAAACGCGCAAGCTGATTGCTTACCCTGCTGTT 1860
Db 1801 CGGTTTACCTGGGATGCGCTAGCTTACAAACGCGCAAGCTGATTGCTTACCCTGCTGTT 1860

Qy 1861 GAAGCGTTATCGTGTGATTATTAACAAAGATCTGCTGCCGAAACCGCCAAAAACCTGGGAA 1920

```
Db 1861 GNAAGGTTATCGCTGATTTATACAAAGATCTGTCGGAACCCGCCAAAAACCTTGGGA 1920
Qy 1921 GAGATCCCGGCGCTGGATAAGAACTGAAGCGAAAGTGAAGAGCGCGCTGATTTCAAC 1980
Db 1921 GAGATCCCGGCGCTGGATAAGAACTGAAGCGAAAGTGAAGAGCGCGCTGATTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATGCTGCTGAGCGGGGTATGCGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTACCTGGCGCTGATGCTGCTGAGCGGGGTATGCGTTCAAG 2040
Qy 2041 TATGAAAACCGCAAGTACGACATTAAGACGTTGGCGGTGATAACGCTGGCGCGAAAGCG 2100
Db 2041 TATGAAAACCGCAAGTACGACATTAAGACGTTGGCGGTGATAACGCTGGCGCGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTCTGTTGACCTGATTAATAAACAACACATGAATGACAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTCTGTTGACCTGATTAATAAACAACACATGAATGACAGACACCGATTAC 2160
Qy 2161 TCCATCGCAGAGCTCCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Db 2161 TCCATCGCAGAGCTCCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCCCGTGG 2220
Qy 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTTATGTTGAACGCTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTTATGTTGAACGCTACTGCGGACCTTC 2280
Qy 2281 AAGGGTCAACCATCAACCGTTCTGCTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Db 2281 AAGGGTCAACCATCAACCGTTCTGCTGGCGTGTGAGCGCAGGTATTAACGCCGCCAGT 2340
Qy 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGGAAACTATCTGCTGACTGATGAAGGTCTG 2400
Db 2341 CCGAACAAGAGCTGGCAAAAGAGTTCTCGGAAACTATCTGCTGACTGATGAAGGTCTG 2400
Qy 2401 GAAAGGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACAGCAAGAG 2460
Db 2401 GAAAGGGTTAATAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACAGCAAGAG 2460
Qy 2461 TTGGCGAAAGATCCACGATTTGCGGCCACCATGGAAGAACGCCAGAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCACGATTTGCGGCCACCATGGAAGAACGCCAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACAATCCCGCAGATGTCGCTTTCTGTTATGCGGTGCTGATGCGGTGATCAACGCC 2580
Db 2521 CCGAACAATCCCGCAGATGTCGCTTTCTGTTATGCGGTGCTGATGCGGTGATCAACGCC 2580
Qy 2581 GCCAGGGTCTGACAGCTGTCGATGAAGCCCTGGAAGAGCGCGACACTAATTCGAGCTCG 2640
Db 2581 GCCAGGGTCTGACAGCTGTCGATGAAGCCCTGGAAGAGCGCGACACTAATTCGAGCTCG 2640
Qy 2641 AACACAACAACAATAACAATAACAATAACAATAACAATAACAATAACAATAACAATAACA 2700
Db 2641 AACACAACAACAATAACAATAACAATAACAATAACAATAACAATAACAATAACAATAACA 2700
```

RESULT 12

```
US-10-263-153-35
; Sequence 35, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.O1
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 35
; LENGTH: 7322
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3399)
; OTHER INFORMATION: pMBP-c2X-Toxop30del10 (52-284aa)
US-10-263-153-35
```

Query Match

Best Local Similarity 81.5%; Score 2690.4; DB 18; Length 7322;

Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1 CCGACACCATCGAATGGTGTCAAAAACCTTTTCGGGTATGCGCATGATAGCCCGCGGAAGA 60

Db 1 CCGACACCATCGAATGGTGTCAAAAACCTTTTCGGGTATGCGCATGATAGCCCGCGGAAGA 60

Qy 61 GTCATTCAGGGTGGTGAATGTAACCAAGTAAAGTTATACGATGTCGACAGTATGCG 120

Db 61 GTCATTCAGGGTGGTGAATGTAACCAAGTAAAGTTATACGATGTCGACAGTATGCG 120

Qy 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACCAAGCCAGCCAGCCAGCTTTCTGCAAAA 180

Db 121 GTGTCTCTTATCAGACCGTTTCCCGCGTGTGAACCAAGCCAGCCAGCCAGCTTTCTGCAAAA 180

Qy 181 CGCGGAAAAAGTGGAGCGCGATGCGGAGCTGAATTAATTCCTCCCAACCGCGTGGCAC 240

Db 181 CGCGGAAAAAGTGGAGCGCGATGCGGAGCTGAATTAATTCCTCCCAACCGCGTGGCAC 240

Qy 241 AACAACTGGCGGCAACAGTCGTGATGTTGGGTTGCCACCTCCAGTCTGGCCCTGC 300

Db 241 AACAACTGGCGGCAACAGTCGTGATGTTGGGTTGCCACCTCCAGTCTGGCCCTGC 300

Qy 301 ACGCGCGTCCGCAAAATGTCGCGCGGATTAATCTCGCGCGCATCAACTGGGTGCGACGG 360

Db 301 ACGCGCGTCCGCAAAATGTCGCGCGGATTAATCTCGCGCGCATCAACTGGGTGCGACGG 360

Qy 361 TGGTGGTTCGATGGTGAACGAAGCGCGCTCGAAGCTGTAAAGCGCGGTGCACAATC 420

Db 361 TGGTGGTTCGATGGTGAACGAAGCGCGCTCGAAGCTGTAAAGCGCGGTGCACAATC 420

Qy 421 TTCTCGCGCAACCGGTGATGCGGTGATCAATTAATCTCGCGCGCATCAACTGGGTGCGACGG 480

Db 421 TTCTCGCGCAACCGGTGATGCGGTGATCAATTAATCTCGCGCGCATCAACTGGGTGCGACGG 480

Qy 481 TTGTGTGGAAGCTGCCCTGCACTAATGTTCCGGGCTTATTTCTTGATGCTCTGACCCAGA 540

Db 481 TTGTGTGGAAGCTGCCCTGCACTAATGTTCCGGGCTTATTTCTTGATGCTCTGACCCAGA 540

Qy 541 CACCCATCAACAGTATTTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGGTGGAGCATC 600

Db 541 CACCCATCAACAGTATTTATTTTCTCCCATGAAGACGGTACGCGACTGGGCGGTGGAGCATC 600

Qy 601 TGGTGGCATTTGGGTACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCGTGTCGG 660

Db 601 TGGTGGCATTTGGGTACAGCAAAATCGCGCTGTTAGCGGGGCCATTAAGTTCGTGTCGG 660

Qy 661 CGCGTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720

Db 661 CGCGTCTGCGTCTGGCTGGCATTAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720

Qy 721 CGGAACGGGAAGCGGACTGGAGTGCCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780

Db 721 CGGAACGGGAAGCGGACTGGAGTGCCATGTCGGTTTTTCAACAAACCATGCAAAATGCTGA 780

Qy 781 ATGAGGGCATGTTCCCACTCGCATGCTGGTTGCGCAACGATCAGATGGCGGTGGCGCAA 840

Db 781 ATGAGGGCATGTTCCCACTCGCATGCTGGTTGCGCAACGATCAGATGGCGGTGGCGCAA 840

Qy 841 TGGCGCCATTAACGAGTCCGGGCTCGCGTGGTGGCGGATATCTCGTACTGAGGATAG 900

Db 841 TGGCGCCATTAACGAGTCCGGGCTCGCGTGGTGGCGGATATCTCGTACTGAGGATAG 900

QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAAGGATTTTC 960
DB 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAAACCAACCATCAAAAGGATTTTC 960
QY 961 GCTCTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
DB 961 GCTCTGCTGGGCAAAACAGCGTGGACCGCTTGTGCAACTCTCTCAGGCGCCAGGCGGTGA 1020
QY 1021 AGGCAATCAGCTGTGTCGCGCTCATCTGCTGAAAGAAACCAACCCCTGGCGGCCAATA 1080
DB 1021 AGGCAATCAGCTGTGTCGCGCTCATCTGCTGAAAGAAACCAACCCCTGGCGGCCAATA 1080
QY 1081 CCGCAACCGCCTCTCCCGCGGTTGGCGAATTCATTAAATGCACTGGCAGCAGAGGTTT 1140
DB 1081 CCGCAACCGCCTCTCCCGCGGTTGGCGAATTCATTAAATGCACTGGCAGCAGAGGTTT 1140
QY 1141 CCGCACTGGAAGCGGCGAGTGACGCGCAACGCAATTAATGTAAGTTAGCTCACTCAATTAG 1200
DB 1141 CCGCACTGGAAGCGGCGAGTGACGCGCAACGCAATTAATGTAAGTTAGCTCACTCAATTAG 1200
QY 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGCG 1260
DB 1201 GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGCACGGTGCACCAATGCTTCTGCG 1260
QY 1261 TCAGGACGCGATCGGAAGCTGTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1320
DB 1261 TCAGGACGCGATCGGAAGCTGTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1320
QY 1321 TGTCTCTCAAGCGCACTCCCGTCTCGATTAATGTTTTTTCGCGGACATCATACCGTT 1380
DB 1321 TGTCTCTCAAGCGCACTCCCGTCTCGATTAATGTTTTTTCGCGGACATCATACCGTT 1380
QY 1381 CTGGCAATATTTCTGAATGAGCTGTGCAATTAATCATCGGCTCGTAAATGTTGGA 1440
DB 1381 CTGGCAATATTTCTGAATGAGCTGTGCAATTAATCATCGGCTCGTAAATGTTGGA 1440
QY 1441 ATTGTAGCGGATTAACATTTTACACAGGAAACGCGTTCGTTTGTGTTTTCAG 1500
DB 1441 ATTGTAGCGGATTAACATTTTACACAGGAAACGCGTTCGTTTGTGTTTTCAG 1500
QY 1501 GCATTTTCAACCAAGGACCATAGATTATGAAACTGAAGAGTAACTGGTAACTCTG 1560
DB 1501 GCATTTTCAACCAAGGACCATAGATTATGAAACTGAAGAGTAACTGGTAACTCTG 1560
QY 1561 ATTAACGCGGATAAGGCTATAACGCTCTGCTGAACTCGGTAAGAAATTCGAGAAAGAT 1620
DB 1561 ATTAACGCGGATAAGGCTATAACGCTCTGCTGAACTCGGTAAGAAATTCGAGAAAGAT 1620
QY 1621 ACCGGAATTAAGTACCGTTGACATCCGGATAACTGGNAGAGAAATTCACACAGTT 1680
DB 1621 ACCGGAATTAAGTACCGTTGACATCCGGATAACTGGNAGAGAAATTCACACAGTT 1680
QY 1681 CGCGCAACTGCGGATGCGCTGACATTAATCTTCTGGGCAACGACCGCTTTGGTGGCTAC 1740
DB 1681 CGCGCAACTGCGGATGCGCTGACATTAATCTTCTGGGCAACGACCGCTTTGGTGGCTAC 1740
QY 1741 GCTCAATCTGCGCTTGTGCTGAATCAACCCCGGACAAAGGTTTCCAGGCAAGCTGTAT 1800
DB 1741 GCTCAATCTGCGCTTGTGCTGAATCAACCCCGGACAAAGGTTTCCAGGCAAGCTGTAT 1800
QY 1801 CCGTTTACCTGGGATGCGTACGTTTCAACGCGAAGCTGATGCTTACCGGATCGCTGTT 1860
DB 1801 CCGTTTACCTGGGATGCGTACGTTTCAACGCGAAGCTGATGCTTACCGGATCGCTGTT 1860
QY 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCGCAACCGCCCAAAACCTGGGA 1920
DB 1861 GAAGCGTTATCGCTGATTATTAACAAAGATCTGCTGCGCAACCGCCCAAAACCTGGGA 1920
QY 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAGGTAAGCGCGCTGATGTTCAAC 1980
DB 1921 GAGATCCCGCGCTGGATTAAGAACTGAAGCGAAGGTAAGCGCGCTGATGTTCAAC 1980

QY 1981 CTGCAAGAACGTAATTCACCTGCGCCCTGATTTGCTGCTGACGGGGTATGCTTCAAG 2040
DB 1981 CTGCAAGAACGTAATTCACCTGCGCCCTGATTTGCTGCTGACGGGGTATGCTTCAAG 2040
QY 2041 TATGAAAAACGCAAGTACGACATTAAGAGACGTTGGCGTGGATACGCTGGCGCAAGCG 2100
DB 2041 TATGAAAAACGCAAGTACGACATTAAGAGACGTTGGCGTGGATACGCTGGCGCAAGCG 2100
QY 2101 GGTCTGACCTTCTCTGTTGACCTGATTAATAAAACAAACACATGAATGACAGACCCGATTAC 2160
DB 2101 GGTCTGACCTTCTCTGTTGACCTGATTAATAAAACAAACACATGAATGACAGACCCGATTAC 2160
QY 2161 TCCATCCGAGAAGCTGCTTTTAAAGGCGAACACAGCGATGACCATCAACGGCCCGTGG 2220
DB 2161 TCCATCCGAGAAGCTGCTTTTAAAGGCGAACACAGCGATGACCATCAACGGCCCGTGG 2220
QY 2221 GCATGGTCCAACTCGACACACCAAGTGAATTTATGGTGTAAACGGTACTTGCACACCTTC 2280
DB 2221 GCATGGTCCAACTCGACACACCAAGTGAATTTATGGTGTAAACGGTACTTGCACACCTTC 2280
QY 2281 AAGGTTCAACCATCCAAACCGTTTCTGTTGGCGTGTGAGCGCAGGTATTAACGGCCAGT 2340
DB 2281 AAGGTTCAACCATCCAAACCGTTTCTGTTGGCGTGTGAGCGCAGGTATTAACGGCCAGT 2340
QY 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTTCTCGAAAACTATCTGCTGATGATGAGGTTCTG 2400
DB 2341 CCGAAACAAAGAGCTGGCAAAAGAGTTTCTCGAAAACTATCTGCTGATGATGAGGTTCTG 2400
QY 2401 GAAGCGTTTAAAGAGCAAAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
DB 2401 GAAGCGTTTAAAGAGCAAAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2460
QY 2461 TTGCGGAAAGATCCAGTATTTGCGCGCAACCATCGGTAAGTCTTACGAGGAGAG 2520
DB 2461 TTGCGGAAAGATCCAGTATTTGCGCGCAACCATCGGTAAGTCTTACGAGGAGAG 2520
QY 2521 CCGAACATCCGAGATGTCGCTTTCTGCTATGCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2580
DB 2521 CCGAACATCCGAGATGTCGCTTTCTGCTATGCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2580
QY 2581 GCCAGCGTCTGAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
DB 2581 GCCAGCGTCTGAGACTGTCGATGAAGCCCTGAAAGACGCGCAGACTAATTCGAGCTCG 2640
QY 2641 AACCAACAAACAAATTAACAAATTAACAAATTAACAAATTAACAAATTAACAAATTAACAA 2700
DB 2641 AACCAACAAACAAATTAACAAATTAACAAATTAACAAATTAACAAATTAACAAATTAAC 2700

RESULT 13

US-10-263-153-25
; Sequence 25, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.O1
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 7352
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3429)

; OTHER INFORMATION: pMBP-c2X-ToxoP30del14C (52-294aa) US-10-263-153-25									
Query Match 81.5%; Score 2690.4; DB 18; Length 7352; Best Local Similarity 99.8%; Pred. No. 0; Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;									
QY	1	CGGACACCATCGAATGGTGGCAAAACCTTTTCGGGGTATGCGCATGATAGGCCCGCGGAAGAGA	60						
DB	1	CGGACACCATCGAATGGTGGCAAAACCTTTTCGGGGTATGCGCATGATAGGCCCGCGGAAGAGA	60						
QY	61	GTCAATTACAGGTGGTGAATGTGAACACAGTAACAGTTATACGATCTCCAGAGTATGCCG	120						
DB	61	GTCAATTACAGGTGGTGAATGTGAACACAGTAACAGTTATACGATCTCCAGAGTATGCCG	120						
QY	121	GTGTCCTTATCAGACCGTTTCCTCCGCTGGTGAACACGCGCACGCCACGCTTTCTCGGAATA	180						
DB	121	GTGTCCTTATCAGACCGTTTCCTCCGCTGGTGAACACGCGCACGCCACGCTTTCTCGGAATA	180						
QY	181	CGCGGAAAGTGAAGCGGGGATGGCGAGCTGAAATTAACATTCCTCAACCGCGTGGCAC	240						
DB	181	CGCGGAAAGTGAAGCGGGGATGGCGAGCTGAAATTAACATTCCTCAACCGCGTGGCAC	240						
QY	241	AACAACTGCGGGCAACAGCTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300						
DB	241	AACAACTGCGGGCAACAGCTGTGCTGATTTGGCGTTGCCACCTCCAGTCTGGCCCTGC	300						
QY	301	ACGCGCGCTCGCAAAATTTGTCGGGGGATTAATCTCGGCGCGATCAACTGGGTGCCAGCG	360						
DB	301	ACGCGCGCTCGCAAAATTTGTCGGGGGATTAATCTCGGCGCGATCAACTGGGTGCCAGCG	360						
QY	361	TGGTGGTGTGATGTAGAAAGAAAGCGGGTGCAGAGCTGTAAAGCGGGTGCACAATC	420						
DB	361	TGGTGGTGTGATGTAGAAAGAAAGCGGGTGCAGAGCTGTAAAGCGGGTGCACAATC	420						
QY	421	TTCTCGCGCAAGCTGAGTGGGCTGATTAATTAATCTCCGCTGGATGACCAAGATGCCA	480						
DB	421	TTCTCGCGCAAGCTGAGTGGGCTGATTAATTAATCTCCGCTGGATGACCAAGATGCCA	480						
QY	481	TTGCTGTGGAAAGCTGCCGTGCAATAATGTTCCGGCGTTAATTTCTGATGCTCTGACCCAGA	540						
DB	481	TTGCTGTGGAAAGCTGCCGTGCAATAATGTTCCGGCGTTAATTTCTGATGCTCTGACCCAGA	540						
QY	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600						
DB	541	CACCCATCAACAGTATTAATTTCTCCATGAAGACGGTACGCGACTGGGCGTGGAGCATC	600						
QY	601	TGGTGCATTTGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG	660						
DB	601	TGGTGCATTTGGTCAACAGCAAAATCGCGCTGTTAGCGGGCCCATTAAGTTCTGCTCGG	660						
QY	661	CGCGTCTGCTGCTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
DB	661	CGCGTCTGCTGCTGGCTGGCAATAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720						
QY	721	CGGAAACGGGAAGCGCATGTGAGTGCATGTCCGGTTTTTCAACAAACCAATGCAAAATGCTGA	780						
DB	721	CGGAAACGGGAAGCGCATGTGAGTGCATGTCCGGTTTTTCAACAAACCAATGCAAAATGCTGA	780						
QY	781	ATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGGCGCA	840						
DB	781	ATGAGGGCATCGTTCCCACTGCGATGCTGGTTGCCAACGATCAGATGGCGCTGGGGCGCA	840						
QY	841	TGCGGCGCATTAACGAGTCCGGGCTGGCGTTGGTGGATATCTCGGTAGTGGGATAGC	900						
DB	841	TGCGGCGCATTAACGAGTCCGGGCTGGCGTTGGTGGGATATCTCGGTAGTGGGATAGC	900						
QY	901	ACGATACCGAAGACAGCTCATGTATATCCCGCGGTTAACACCAATCAAAACAGGATTTTC	960						
DB	901	ACGATACCGAAGACAGCTCATGTATATCCCGCGGTTAACACCAATCAAAACAGGATTTTC	960						
QY	961	GCCTGCTGGGGCAACACGAGCTGGACCGCTTGTGCAATCTCTCAGGCGCCAGGCGGTGA	1020						
DB	961	GCCTGCTGGGGCAACACGAGCTGGACCGCTTGTGCAATCTCTCAGGCGCCAGGCGGTGA	1020						

DB	961	GCCTGCTGGGGCAACACGAGCGTGGACCGCTTGCTGCAACTCTCTCAGGCGCCAGGCGGTGA	1020						
QY	1021	AGGCAATACAGCTGTGTCCTCCCTCTCACTGGTGAAGAAAAACCACTTCGCGCCCAATA	1080						
DB	1021	AGGCAATACAGCTGTGTCCTCCCTCTCACTGGTGAAGAAAAACCACTTCGCGCCCAATA	1080						
QY	1081	CGCAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAATCGAGCTGGCAGCAGAGTTT	1140						
DB	1081	CGCAACCGCTCTCTCCCGCGGTTGGCGGATTCATTAATCGAGCTGGCAGCAGAGTTT	1140						
QY	1141	CCGACTCGAAAGCGGGCAGTGAGCGCAACCAATTAATGTAGTTAGTCTACTCATTTAG	1200						
DB	1141	CCGACTCGAAAGCGGGCAGTGAGCGCAACCAATTAATGTAGTTAGTCTACTCATTTAG	1200						
QY	1201	GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGACGGTGCACCAATGCTTCTGCG	1260						
DB	1201	GCACAATTTCTCATGTTTGACAGCTTATCATCGACTGACGGTGCACCAATGCTTCTGCG	1260						
QY	1261	TCAGGACGCGATCGGAAGCTGTGATGCTGTGCGAGTCTGTAATCACTGCATTAATTCG	1320						
DB	1261	TCAGGACGCGATCGGAAGCTGTGATGCTGTGCGAGTCTGTAATCACTGCATTAATTCG	1320						
QY	1321	TGTGCTCAAGCGCAGCTCCCGTTCTGATTAATGTTTTTGGCGGCAATCAATAACGGTT	1380						
DB	1321	TGTGCTCAAGCGCAGCTCCCGTTCTGATTAATGTTTTTGGCGGCAATCAATAACGGTT	1380						
QY	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA	1440						
DB	1381	CTGGCAAAATTTCTGAAATGAGCTGTTGACAAATTAATCATCGGCTCGTATATGTTGGA	1440						
QY	1441	ATTGTGAGCGGATTAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTTTTCA	1500						
DB	1441	ATTGTGAGCGGATTAACAATTTTCAACAGGAAACAGCCAGTCCGTTTAGTGTTTTTCA	1500						
QY	1501	GCACCTTCAACCAAGGACCATAGATTATGAAATCTGAAGAGGTAATACTGTAATCTGG	1560						
DB	1501	GCACCTTCAACCAAGGACCATAGATTATGAAATCTGAAGAGGTAATACTGTAATCTGG	1560						
QY	1561	ATTAACCGCGATAAAGGCTATAACCGGCTCGCTGAAGTCCGTAAGAAATTCGAGAAAGAT	1620						
DB	1561	ATTAACCGCGATAAAGGCTATAACCGGCTCGCTGAAGTCCGTAAGAAATTCGAGAAAGAT	1620						
QY	1621	ACCGGAATTAAGTACCGGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT	1680						
DB	1621	ACCGGAATTAAGTACCGGTTGAGCATCCGATTAACCTGGAAGAGAAATTCACAGGTT	1680						
QY	1681	GGGCAACTGCGCATGCGCCCTGACATTAATCTTCTGGGCACACGCGCTTCTGGTGCTAC	1740						
DB	1681	GGGCAACTGCGCATGCGCCCTGACATTAATCTTCTGGGCACACGCGCTTCTGGTGCTAC	1740						
QY	1741	GCTCAATCTGGCTGTGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT	1800						
DB	1741	GCTCAATCTGGCTGTGCTGAAATCAACCCCGGACAAAGCGTTCCAGGACAAGCTGTAT	1800						
QY	1801	CGGTTTACCTGGATGCGGTACGTTTAAACGGAAGCTGATTTGCTTACCCGATCGCTGTT	1860						
DB	1801	CGGTTTACCTGGATGCGGTACGTTTAAACGGAAGCTGATTTGCTTACCCGATCGCTGTT	1860						
QY	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA	1920						
DB	1861	GAAGCGTTATCGCTGATTTATAACAAAGATCTGCTGCCGAAACCCGCCAAAACCTGGGAA	1920						
QY	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGGTGAAGAGCGCGCTGATGTTTCAAC	1980						
DB	1921	GAGATCCCGCGCTGGATAAAGAACTGAAAGCGAAGGTGAAGAGCGCGCTGATGTTTCAAC	1980						
QY	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATTTGCTGTGACGGGGTTATTCGCTTCAAG	2040						
DB	1981	CTGCAAGAACCGTACTTCACTGGCGCTGATTTGCTGTGACGGGGTTATTCGCTTCAAG	2040						
QY	2041	TATGAAAACGCGAAGTACGACATTAAGACGTGGCGGTGGATTAACGCTGGCGGGAAGCG	2100						
DB	2041	TATGAAAACGCGAAGTACGACATTAAGACGTGGCGGTGGATTAACGCTGGCGGGAAGCG	2100						

```
QY 2101 GGTCTGACCTTCTGTTGACCTGATTAAAAACAACACATGAATGCAGACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGTTGACCTGATTAAAAACAACACATGAATGCAGACACCGATTAC 2160
QY 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAAACAGCGATGACCATCAACGGCCGTTGG 2220
Db 2161 TCCATCGCAGAGCTGCTTTAATAAGCGGAAAACAGCGATGACCATCAACGGCCGTTGG 2220
QY 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTATGGTGAACGGTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAATCGACACAGCAAGTGAATTATGGTGAACGGTACTGCGGACCTTC 2280
QY 2281 AAGGFTCAACCATCCAAACCGTTGCTGGCGTGTGAGCGCAGGTATTAAACGCCGCCAGT 2340
Db 2281 AAGGFTCAACCATCCAAACCGTTGCTGGCGTGTGAGCGCAGGTATTAAACGCCGCCAGT 2340
QY 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGATGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCCTCGAAAACTATCTGCTGATGATGAAGTCTG 2400
QY 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
Db 2401 GAAGCGGTTAATAAAGACAAACCGCTGGGTGCGGTAGCGCTGAAGTCTTACGAGGAAG 2460
QY 2461 TTGGCGAAAGATCCAGTATTGCCGCCACCATGGAAAAACGCCAGAAAAGGTGAATCATG 2520
Db 2461 TTGGCGAAAGATCCAGTATTGCCGCCACCATGGAAAAACGCCAGAAAAGGTGAATCATG 2520
QY 2521 CCGAACATCCCGCAGATGCTCGCTTCTGGTATGCGGTGCGTACTGCGGTGATCAAGCC 2580
Db 2521 CCGAACATCCCGCAGATGCTCGCTTCTGGTATGCGGTGCGTACTGCGGTGATCAAGCC 2580
QY 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
Db 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTCG 2640
QY 2641 AACACACACAAATTAACAATAACAACTCGGGATCGAGGAAAGATTCAGAATTC 2700
Db 2641 AACACACACAAATTAACAATAACAACTCGGGATCGAGGAAAGATTCAGAATTC 2700
```

RESULT 14

```
US-10-263-153-20
; Sequence 20, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Gineburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN.
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263.153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 7370
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3447)
; OTHER INFORMATION: pMBP-c2X-Toxo30del3C (52-300aa)
US-10-263-153-20
```

Query Match 81.5%; Score 2690.4; DB 18; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
QY 1 CGGACACCATCGAATGGTGGAAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
Db 1 CGGACACCATCGAATGGTGGAAAAACCTTTCCGGGTATGGCATGATAGCGCCCGGAAGAGA 60
QY 61 GTCAAATTCAGGGTGGTGAATGTGAACACAGTAAACGTTATACGATGTCGAGAGTATCGCG 120
Db 61 GTCAAATTCAGGGTGGTGAATGTGAACACAGTAAACGTTATACGATGTCGAGAGTATCGCG 120
QY 121 GTGTCTCTTATCAGACCGTTTCCCGTGTGGTGAACACAGGCGAGCCAGTTCCTCGAAAA 180
Db 121 GTGTCTCTTATCAGACCGTTTCCCGTGTGGTGAACACAGGCGAGCCAGTTCCTCGAAAA 180
QY 181 CGCGGGAAGTGGAGCGCGATGCGGAGCTGAATTAATTCCTCAACCGCGTGGCAC 240
Db 181 CGCGGGAAGTGGAGCGCGATGCGGAGCTGAATTAATTCCTCAACCGCGTGGCAC 240
QY 241 AACAACTGGGGGGCAACAGTTCGTTGCTGATTTGGCGTTGCGACCTCCAGTCTGGCCCTGC 300
Db 241 AACAACTGGGGGGCAACAGTTCGTTGCTGATTTGGCGTTGCGACCTCCAGTCTGGCCCTGC 300
QY 301 ACGCCCGTGCAGAAATTTGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
Db 301 ACGCCCGTGCAGAAATTTGTCGCGCGATTAATCTCGCGCGATCAACTGGGTGCCAGCG 360
QY 361 TGGTGGTGTGATGTAAGAACGAGCGCGTGAAGCCTGTAAGCGCGCGGTGCACAAATC 420
Db 361 TGGTGGTGTGATGTAAGAACGAGCGCGTGAAGCCTGTAAGCGCGCGGTGCACAAATC 420
QY 421 TTCTCGCGCAACCGTCAAGTGGCTGATCATTAATCTATCGCTGGATGACAGAGTGCCCA 480
Db 421 TTCTCGCGCAACCGTCAAGTGGCTGATCATTAATCTATCGCTGGATGACAGAGTGCCCA 480
QY 481 TTCTGTGGAGAGTGGCTGCACTAAATTTTCGCGCGTTATTTCTTGATGTCTCTGACCAGA 540
Db 481 TTCTGTGGAGAGTGGCTGCACTAAATTTTCGCGCGTTATTTCTTGATGTCTCTGACCAGA 540
QY 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACCGTACGCGACTGGGCGTGGAGCATC 600
Db 541 CACCCATCAACAGTATTAATTTCTCCATGAAGACCGTACGCGACTGGGCGTGGAGCATC 600
QY 601 TGGTGGATTTGGTCAACAGCAAAATCGCGTGTAGCGGGCCCATTAAGTTCGTCTCGG 660
Db 601 TGGTGGATTTGGTCAACAGCAAAATCGCGTGTAGCGGGCCCATTAAGTTCGTCTCGG 660
QY 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
Db 661 CGCGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAG 720
QY 721 CGGAACGGGAAGCGACTGGAGTGCCATGTCCGGTTTTCACAAACCATGCAAAATGCTGA 780
Db 721 CGGAACGGGAAGCGACTGGAGTGCCATGTCCGGTTTTCACAAACCATGCAAAATGCTGA 780
QY 781 ATGAGGGCATGCTTCCACTGCGATGCTGGTTCGCAACGATCAGATGGCGCTGGGCGCAA 840
Db 781 ATGAGGGCATGCTTCCACTGCGATGCTGGTTCGCAACGATCAGATGGCGCTGGGCGCAA 840
QY 841 TGGCGCCATTAACCGAGTCCGGCTGGCGTGGTGGCGGATATCTCGGTAGTGGGATAG 900
Db 841 TGGCGCCATTAACCGAGTCCGGCTGGCGTGGTGGCGGATATCTCGGTAGTGGGATAG 900
QY 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAAAAGGATTTTC 960
Db 901 ACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAACCAACCATCAAAAGGATTTTC 960
QY 961 GCCTGCTGGGCAAAACAGCGTGGAGCGGTGCTGCTCAACTCTCAGGGCCAGGCGGTGA 1020
Db 961 GCCTGCTGGGCAAAACAGCGTGGAGCGGTGCTGCTCAACTCTCAGGGCCAGGCGGTGA 1020
QY 1021 AGGGCAATCAGTGTTCGCCGCTCTCACTGGTGAAGAAAAAACCAACCTTGGCGCCCAATA 1080
Db 1021 AGGGCAATCAGTGTTCGCCGCTCTCACTGGTGAAGAAAAAACCAACCTTGGCGCCCAATA 1080
QY 1081 CGCAAAACCGCCTCTCCCGCGGTTGGCGGATTTCAATTAATGAGCTGGCAACGACAGTTT 1140
```

Db 1081 ||||| CGAACCCTCTCCCGCGGTTGGCGATTCAATTAATGAGTGGCAAGAGTTT 1140
Qy 1141 CCCGACTGGAAGCGGCGAGTGAGCGAACCAATTAATGTAGTTAGCTCACTCAATTAG 1200
Db 1141 CCCGACTGGAAGCGGCGAGTGAGCGAACCAATTAATGTAGTTAGCTCACTCAATTAG 1200
Qy 1201 GCACAAATCTCATGTTTGACAGCTTATCATGACTGCGAGGTGACCAATGCTTCTGGCG 1260
Db 1201 GCACAAATCTCATGTTTGACAGCTTATCATGACTGCGAGGTGACCAATGCTTCTGGCG 1260
Qy 1261 TCAGCAGCATCGAAGCTGTGGTATGCTGTGCAGGTGTAATCACTGATCAATTCG 1320
Db 1261 TCAGCAGCATCGAAGCTGTGGTATGCTGTGCAGGTGTAATCACTGATCAATTCG 1320
Qy 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTGGCGGACATCAATACGGTT 1380
Db 1321 TGTGCTCAAGCGCACTCCCGTTCTGGATAATGTTTTTGGCGGACATCAATACGGTT 1380
Qy 1381 CTGGCAATATCTGAATAGCTGTTGACAAATTAATCATGCGCTGTAATGTGTGA 1440
Db 1381 CTGGCAATATCTGAATAGCTGTTGACAAATTAATCATGCGCTGTAATGTGTGA 1440
Qy 1441 ATTGTAGCGGATAACAATTTACACAGAGAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Db 1441 ATTGTAGCGGATAACAATTTACACAGAGAACAGCCAGTCCGTTTAGTGTTCACGA 1500
Qy 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAATCTGG 1560
Db 1501 GCACCTTCAACCAAGGACCATAGATTATGAAAACTGAAGAGGTAAACTGGTAATCTGG 1560
Qy 1561 ATTAACGGCGATAAAGGCTATACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Db 1561 ATTAACGGCGATAAAGGCTATACCGTCTCGCTGAAGTCGGTAAGAAATTCAGAAAGAT 1620
Qy 1621 ACCGAAATTAAGTCACCGTTGAGCATCCGATTAACCTGGAGAGAAATTCACAGGTT 1680
Db 1621 ACCGAAATTAAGTCACCGTTGAGCATCCGATTAACCTGGAGAGAAATTCACAGGTT 1680
Qy 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTGGGCACACGCGCTTTGGTGGCTAC 1740
Db 1681 GCGCAACTGGCGATGGCCCTGACATTAATCTTGGGCACACGCGCTTTGGTGGCTAC 1740
Qy 1741 GCTCAATCTGGCTGTTGGCTGAATCAACCCGGAACAAGCGTTTCAGACAAAGCTGTAT 1800
Db 1741 GCTCAATCTGGCTGTTGGCTGAATCAACCCGGAACAAGCGTTTCAGACAAAGCTGTAT 1800
Qy 1801 CCGTTTACCTGGATGCCGTACGTTTACACGCAAGCTGATGCTTACCGATCGCTGTT 1860
Db 1801 CCGTTTACCTGGATGCCGTACGTTTACACGCAAGCTGATGCTTACCGATCGCTGTT 1860
Qy 1861 GAAGGTTATCGCTGATTTATAACAAAGATCTGTCGCGAACCCGCCAAAACCTGGGAA 1920
Db 1861 GAAGGTTATCGCTGATTTATAACAAAGATCTGTCGCGAACCCGCCAAAACCTGGGAA 1920
Qy 1921 GAGATCCCGGCTCGATAAGAACTGAAGCGAAAGTGAAGCGCGCTGATGTTCAAC 1980
Db 1921 GAGATCCCGGCTCGATAAGAACTGAAGCGAAAGTGAAGCGCGCTGATGTTCAAC 1980
Qy 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGACGCGGGGTTATCGGTTCAAG 2040
Db 1981 CTGCAAGAACCGTACTTCACTGGCGCTGATGCTGACGCGGGGTTATCGGTTCAAG 2040
Qy 2041 TATGAAACCGCAAGTACGACATTAAGACGTGGCGGTGATAACCTGGCGCGAAAGCG 2100
Db 2041 TATGAAACCGCAAGTACGACATTAAGACGTGGCGGTGATAACCTGGCGCGAAAGCG 2100
Qy 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACATGAATGACACACCGATTAC 2160
Db 2101 GGTCTGACCTTCTGGTTGACCTGATTAATAAACAACATGAATGACACACCGATTAC 2160
Qy 2161 TCCATCGCAAGAGCTGCTTTAATTAAGCGGAAACAGCGATGACCATCAACCGGCCGTGG 2220

Db 2161 TCCATCGCAAGAGCTGCTTTAATAAAGCGAAACAGCATGACCATCAACGCGCCGTGG 2220
Qy 2221 GCATGGTCCAAACATCGACACACAGCAAAAGTGAATTAATGTGTAAACGTTACTGCGGACCTTC 2280
Db 2221 GCATGGTCCAAACATCGACACACAGCAAAAGTGAATTAATGTGTAAACGTTACTGCGGACCTTC 2280
Qy 2281 AAGGTCACACCATCAACACCGTTGTTGGCGTGTGAGCGCAGGTATTAACGCGCCAGT 2340
Db 2281 AAGGTCACACCATCAACACCGTTGTTGGCGTGTGAGCGCAGGTATTAACGCGCCAGT 2340
Qy 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGATGATGAAGTCTG 2400
Db 2341 CCGAACAAAGAGCTGGCAAAAGAGTTCTCGAAAACTATCTGCTGATGATGAAGTCTG 2400
Qy 2401 GAAGCGTTAATAAAGACAAACCGCTGGTGGCGTGTGAGCGTGAAGTCTTACAGAGAAAG 2460
Db 2401 GAAGCGTTAATAAAGACAAACCGCTGGTGGCGTGTGAGCGTGAAGTCTTACAGAGAAAG 2460
Qy 2461 TTGGGAAAGATCCACGTTATTGCGCCACCATGGAACGCCGAGAAAGGTGAATCATG 2520
Db 2461 TTGGGAAAGATCCACGTTATTGCGCCACCATGGAACGCCGAGAAAGGTGAATCATG 2520
Qy 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTGCTGCTGATCAACGCC 2580
Db 2521 CCGAACATCCCGCAGATGTCGCTTTCTGGTATGCGGTGCTGCTGCTGATCAACGCC 2580
Qy 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTG 2640
Db 2581 GCCAGCGTCTGACAGTGTGATGAAGCCCTGAAAGACGCGCAGACTAAATTCGAGCTG 2640
Qy 2641 AACCAACACCAATAACAATAACAACCTCGGATCGAGGAGGATTCAGAAATTC 2700
Db 2641 AACCAACACCAATAACAATAACAACCTCGGATCGAGGAGGATTCAGAAATTC 2700

RESULT 15

US-10-263-153-61
; Sequence 61, Application US/10263153
; Publication No. US20040067239A1
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Maine, Gregory T.
; APPLICANT: Patel, Chandu B.
; APPLICANT: Ginsburg, Sanford R.
; APPLICANT: Bliese, Timothy R.
; TITLE OF INVENTION: GENETICALLY ENGINEERED P30 ANTIGEN,
; TITLE OF INVENTION: IMPROVED ANTIGEN COCKTAIL, AND USES THEREOF
; FILE REFERENCE: 6984.US.01
; CURRENT APPLICATION NUMBER: US/10/263,153
; CURRENT FILING DATE: 2002-10-22
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 61
; LENGTH: 7370
; TYPE: DNA
; ORGANISM: Toxoplasma gondii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1528)...(3447)
; OTHER INFORMATION: pMBP-c2X-Toxop30MIX1
US-10-263-153-61

Query Match 81.5%; Score 2690.4; DB 18; Length 7370;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 2694; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qy 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGGGTATGGCATGATAGCCCGGAAGAGA 60
Db 1 CCGACACCATCGAATGGTGCAAAACCTTTTCGGGTATGGCATGATAGCCCGGAAGAGA 60
Qy 61 GTCAATTCAGGTCGTGAATGTGAAACCAAGTAACGTTATACGATCTCCAGAGTATCGG 120
Db 61 GTCAATTCAGGTCGTGAATGTGAAACCAAGTAACGTTATACGATCTCCAGAGTATCGG 120

Qy	121	GTGTCTCTTATCAGACCGTGTTC	CGCGTGTGTGAACAGGCCAGCCACGTTTCTCGGAAA	180
Db	121	GTGTCTCTTATCAGACCGTGTTC	CGCGTGTGTGAACAGGCCAGCCACGTTTCTCGGAAA	180
Qy	181	CGCGGGAAAAGTGGAAAGCGCGCAT	TGCGGAGCTGAATTACATTCCCAACCGCGTGGCAC	240
Db	181	CGCGGGAAAAGTGGAAAGCGCGCAT	TGCGGAGCTGAATTACATTCCCAACCGCGTGGCAC	240
Qy	241	AACAACTGGCGGGCAAAACAGT	CGTGTCTGATTGCGGTGTCACCTCCAGTCTGGCCCTGC	300
Db	241	AACAACTGGCGGGCAAAACAGT	CGTGTCTGATTGCGGTGTCACCTCCAGTCTGGCCCTGC	300
Qy	301	ACGCGCGTGCMAATTTGTCGCGG	CGAATAAATCTCGCGCGATCACTAAGTGGGTGCCAGCG	360
Db	301	ACGCGCGTGCMAATTTGTCGCGG	CGAATAAATCTCGCGCGATCACTAAGTGGGTGCCAGCG	360
Qy	361	TGCGTGTGTCGATGGTAGAACGA	AGCGCGTGGAGCCTGTAAAGCGCGGTGACAATTC	420
Db	361	TGCGTGTGTCGATGGTAGAACGA	AGCGCGGTGGAGCCTGTAAAGCGCGGTGACAATTC	420
Qy	421	TTCTCTGCGCAAACGCGTCACT	TGGCTGATCAATTAATCTCGCTGGATGACCAAGGATGCCA	480
Db	421	TTCTCTGCGCAAACGCGTCACT	TGGCTGATCAATTAATCTCGCTGGATGACCAAGGATGCCA	480
Qy	481	TTGCTGTGGAAGCTGCCTGCAC	TAAATGTTCCGGCGTTAATCTCGATGCTCTGACCA	540
Db	481	TTGCTGTGGAAGCTGCCTGCAC	TAAATGTTCCGGCGTTAATCTCGATGCTCTGACCA	540
Qy	541	CACCCATCAACAGTATTATTTCT	CCCATGAAGACGGTACGCGATCGGCGGTGGAGCATC	600
Db	541	CACCCATCAACAGTATTATTTCT	CCCATGAAGACGGTACGCGATCGGCGGTGGAGCATC	600
Qy	601	TGGTCGCATTGGGTCACCAAG	CAAAATCGCGCTGTTAGCGGGCCCAATTAAGTCTGTC	660
Db	601	TGGTCGCATTGGGTCACCAAG	CAAAATCGCGCTGTTAGCGGGCCCAATTAAGTCTGTC	660
Qy	661	CGCGTCTCGTCTGGCTGGCA	TAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Db	661	CGCGTCTCGTCTGGCTGGCA	TAAATATCTCACTCGCAATCAAAATTCAGCCGATAG	720
Qy	721	CGGAACGGGAAGGCGATGGA	GTGCGATGTCGCGTTTTCACAAACCATGCAATGCTGA	780
Db	721	CGGAACGGGAAGGCGATGGA	GTGCGATGTCGCGTTTTCACAAACCATGCAATGCTGA	780
Qy	781	ATGAGGGCATCGTTCCCACT	CCGATGCTGGTTCGCAACGATCAGATGCGGTGGCGCA	840
Db	781	ATGAGGGCATCGTTCCCACT	CCGATGCTGGTTCGCAACGATCAGATGCGGTGGCGCA	840
Qy	841	TGCGCGCATTACCGAGTCCGG	GTCCGGTGTGGCGATCTCGGTAGTGGGATACG	900
Db	841	TGCGCGCATTACCGAGTCCGG	GTCCGGTGTGGCGATCTCGGTAGTGGGATACG	900
Qy	901	ACGATACCGAAGACAGCTCAT	GTATTATCCCGCGTTAAACACCATCAACAGGATTTTC	960
Db	901	ACGATACCGAAGACAGCTCAT	GTATTATCCCGCGTTAAACACCATCAACAGGATTTTC	960
Qy	961	GCCTGTGGGGCAAAACAGCG	TGGAACCGTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020
Db	961	GCCTGTGGGGCAAAACAGCG	TGGAACCGTTGTGCAACTCTCTCAGGGCCAGCGGTGA	1020
Qy	1021	AGGGCAATCAGCTGTTGCCG	CTCATCTGGTGAAGAAAACCACTTGGCGCCCAATA	1080
Db	1021	AGGGCAATCAGCTGTTGCCG	CTCATCTGGTGAAGAAAACCACTTGGCGCCCAATA	1080
Qy	1081	CGAAACCGGCTCTCCCGCG	CTTCACTGGTGAAGAAAACCACTTGGCGCCCAATA	1140
Db	1081	CGAAACCGGCTCTCCCGCG	CTTCACTGGTGAAGAAAACCACTTGGCGCCCAATA	1140
Qy	1141	CCCGATCGGAAGCGGGCAGT	GAGCAACGCAATTAATGTGAGTGTAGTCACTCAATTAG	1200
Db	1141	CCCGATCGGAAGCGGGCAGT	GAGCAACGCAATTAATGTGAGTGTAGTCACTCAATTAG	1200

QY	1201	GCAAAATCTCATGTTTGA	CAGCTTATCATCGA	CTGCA	CGGTGCACAAATGCTTCTGGCG	1260		
DB	1201	GCAAAATCTCATGTTTGACAGCTTATCATCGA	CTGCA	CGGTGCACCAATGCTTCTGGCG	1260			
QY	1261	TCAGGCAGCCATCGGAAGCTGTGTA	TGGCTGTG	CAGGT	CGTAAATC	ATTCATGCAATAATTCG	1320	
DB	1261	TCAGGCAGCCATCGGAAGCTGTGTA	TGGCTGTG	CAGGT	CGTAAATC	ATTCATGCAATAATTCG	1320	
QY	1321	TGTGCGCTCAAGCGCA	CTCCGGTCTCGATAA	TGTTTTTT	CGCGCGACATCA	TAAACGGTT	1380	
DB	1321	TGTGCGCTCAAGCGCA	CTCCGGTCTCGATAA	TGTTTTTT	CGCGCGACATCA	TAAACGGTT	1380	
QY	1381	CTGGCAATATCTCAAAATAGCTGTTGA	CAATTAATCAT	CGGCTCGTAAAT	CGTAAATGTTGGA	1440		
DB	1381	CTGGCAATATCTCAAAATAGCTGTTGA	CAATTAATCAT	CGGCTCGTAAAT	CGTAAATGTTGGA	1440		
QY	1441	ATTGTGAGCGGATAACAAATTTCA	CAGGAAACAGCCAGT	CCGTTT	AGGTGTTT	TCACGA	1500	
DB	1441	ATTGTGAGCGGATAACAAATTTCA	CAGGAAACAGCCAGT	CCGTTT	AGGTGTTT	TCACGA	1500	
QY	1501	GCACCTTCA	CCAAAGGACCATAGAT	TATGA	AAATCGAAGAGTAA	ATCTGTAATCTGG	1560	
DB	1501	GCACCTTCA	CCAAAGGACCATAGAT	TATGA	AAATCGAAGAGTAA	ATCTGTAATCTGG	1560	
QY	1561	ATTAAACGCGGATAAAGGCTA	TAAAGGCTTAC	CGGTCTCGCTG	AAATCGGTAAGAAATTCG	GAGAAAGAT	1620	
DB	1561	ATTAAACGCGGATAAAGGCTA	TAAAGGCTTAC	CGGTCTCGCTG	AAATCGGTAAGAAATTCG	GAGAAAGAT	1620	
QY	1621	ACCGGAATTA	AAAGTCA	CCGTGTAG	CACTCGGAT	TAAACTCGGAAGAGAAATTC	1680	
DB	1621	ACCGGAATTA	AAAGTCA	CCGTGTAG	CACTCGGAT	TAAACTCGGAAGAGAAATTC	1680	
QY	1681	GCGGCAACTG	CGCATG	CCCTGAC	ATTCTTCTGGG	CACAGCGCTTTGGTGGCTAC	1740	
DB	1681	GCGGCAACTG	CGCATG	CCCTGAC	ATTCTTCTGGG	CACAGCGCTTTGGTGGCTAC	1740	
QY	1741	GCTCAATCTGG	CCGTGTTGGTGA	AAATCAC	CCCGGACAAAGCGTTC	CCAGCAAGCTGTAT	1800	
DB	1741	GCTCAATCTGG	CCGTGTTGGTGA	AAATCAC	CCCGGACAAAGCGTTC	CCAGCAAGCTGTAT	1800	
QY	1801	CCGTTTACCTGG	GATGCGGTA	CGTTAC	ACGCAAGCTGAT	TGCTTACCGCATCGCTGTT	1860	
DB	1801	CCGTTTACCTGG	GATGCGGTA	CGTTAC	ACGCAAGCTGAT	TGCTTACCGCATCGCTGTT	1860	
QY	1861	GAAGCGTTAT	CGCTGATTTAT	TAAAGAACTG	CTGCTCCGAA	CCCGCAAAACCTGGGA	1920	
DB	1861	GAAGCGTTAT	CGCTGATTTAT	TAAAGAACTG	CTGCTCCGAA	CCCGCAAAACCTGGGA	1920	
QY	1921	GAGATCCGCGCTGATTAAGAACTGA	AGCAAGTAA	GCGGAAGTAA	GAGCGGCTGATTTCAAC	1980		
DB	1921	GAGATCCGCGCTGATTAAGAACTGA	AGCAAGTAA	GCGGAAGTAA	GAGCGGCTGATTTCAAC	1980		
QY	1981	CTGCAAGAAC	CGGTA	CTTCACTTGG	CGGCTGAT	TGCTGCTGACG	GGGTTATGCGTTCAAG	2040
DB	1981	CTGCAAGAAC	CGGTA	CTTCACTTGG	CGGCTGAT	TGCTGCTGACG	GGGTTATGCGTTCAAG	2040
QY	2041	TATGA	AAACGGCAAGTACGA	CTTAAG	ACGTTGGG	CGTGGATAAC	CGTGGCGGAAAGCG	2100
DB	2041	TATGA	AAACGGCAAGTACGA	CTTAAG	ACGTTGGG	CGTGGATAAC	CGTGGCGGAAAGCG	2100
QY	2101	GGTCTG	CACCTTCTGGT	TGACCTG	ATTA	AAACAAACATG	ATGACGACCGATTAC	2160
DB	2101	GGTCTG	CACCTTCTGGT	TGACCTG	ATTA	AAACAAACATG	ATGACGACCGATTAC	2160
QY	2161	TCCATCG	GAGAGCTGCCTTTA	TAAAGGCG	AAACAGCG	ATGACCACTCA	ACGGCCCGTGG	2220
DB	2161	TCCATCG	GAGAGCTGCCTTTA	TAAAGGCG	AAACAGCG	ATGACCACTCA	ACGGCCCGTGG	2220
QY	2221	GATG	GTCCAAATCGA	CCAGCAAGTGA	ATTTATG	TGTAACGGT	ACTGCGACCTTC	2280
DB	2221	GATG	GTCCAAATCGA	CCAGCAAGTGA	ATTTATG	TGTAACGGT	ACTGCGACCTTC	2280
QY	2281	AGGGTCA	ACCATCCAAACCGT	TTCGTTGG	CGTGTG	AGCGCAGGTAT	TAAACCGCGCAGT	2340

Db	2281	AAGGGTCAACCATCCAAACCGTTGCTTGGGTGCTGAGCGCAGGTATTAAACGGCGCCAGT	2340
Qy	2341	CCGAACAAAGAGCTGGCAAAGAGTTCCTCGAAACCTATCTGCTGACTGATGAAGTCTG	2400
Db	2341	CCGAACAAAGAGCTGGCAAAGAGTTCCTCGAAACCTATCTGCTGACTGATGAAGGTCTG	2400
Qy	2401	GAAGCGGTTAATAAAGACAAACCCCTGGGTGCCGTAGCGTGAAGTCTTACGAGGAAAGAG	2460
Db	2401	GAAGCGGTTAATAAAGACAAACCCCTGGGTGCCGTAGCGTGAAGTCTTACGAGGAAAGAG	2460
Qy	2461	TTGCGCAAAGATCCACGCTATTGCGCGCCACCATGGAAACCGCCAGAAAGGTGAAATCATG	2520
Db	2461	TTGCGCAAAGATCCACGCTATTGCGCGCCACCATGGAAACCGCCAGAAAGGTGAAATCATG	2520
Qy	2521	CCGAACATCCCGCAGATGTCCGCTTTCTGGTATGCCGTGCGTACTCGCGTGATCAACGCC	2580
Db	2521	CCGAACATCCCGCAGATGTCCGCTTTCTGGTATGCCGTGCGTACTCGCGTGATCAACGCC	2580
Qy	2581	GCCAGCGGTGCTAGACTGTGCGATGAAGCCCTGAAAGACCGCAGACTAAATTCGAGTCTG	2640
Db	2581	GCCAGCGGTGCTAGACTGTGCGATGAAGCCCTGAAAGACCGCAGACTAAATTCGAGTCTG	2640
Qy	2641	AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTTCAGAAATTC	2700
Db	2641	AACAAACAACAATAACAATAACAACCTCGGGATCGAGGGAAGGATTTCAGAAATTC	2700

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 2, 2005, 10:29:49 ; Search time 6743.2 Seconds
(without alignments)
18627.964 Million cell updates/sec

Title: US-09-765-555B-17
Perfect score: 3300
Sequence: 1 ccgacacatcgatgtgc.....acgacgttcggactacgct 3300

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:
1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	778	23.6	794	5	BQ751936	BQ751936 EST632499
2	770.6	23.4	779	7	CK394523	CK394523 hggada4D0
3	768	23.3	769	7	CK118014	CK118014 218n06.p1
4	762	23.1	793	5	BQ751655	BQ751655 EST632218
5	755.4	22.9	806	5	BQ751220	BQ751220 EST631783
6	684.8	20.8	689	6	CB863541	CB863541 HH04A089
7	676.4	20.5	712	1	AL038548	AL038548 DKF2P566E
8	676	20.5	710	8	AY080106	AY080106 AY080106
9	647.4	19.6	660	1	AL044483	AL044483 DKF2P4341
10	641	19.4	641	4	BJ063872	BJ063872 BJ063872
11	615.8	18.7	624	6	CB863814	CB863814 HH07B20Y
12	612	18.5	794	1	AL045353	AL045353 DKF2P434B
13	603	18.3	752	9	CL655575	CL655575 PRI0124a
14	589.8	17.9	608	8	AY080094	AY080094 AY080094
15	586.4	17.8	589	1	AL037742	AL037742 DKF2P564C
16	581.2	17.6	601	1	AV594538	AV594538 AV594538
17	579	17.5	579	1	AL043868	AL043868 DKF2P434C
18	570	17.3	782	7	CK394246	CK394246 hggada1B1
19	569	17.2	774	1	AL045337	AL045337 DKF2P434A
20	563.2	17.1	585	9	CL658017	CL658017 PRI0130b
21	562.6	17.0	723	1	AL039416	AL039416 DKF2P434L
22	549.4	16.6	753	1	AL045341	AL045341 DKF2P434B
23	546.8	16.6	782	1	AL037051	AL037051 DKF2P564I
24	545.4	16.5	547	6	CA895956	CA895956 B0195F10-

C 25	545	16.5	648	7	CK394349	CK394349 hggada2D0
C 26	540.2	16.4	569	1	AL039077	AL039077 DKF2P566G
C 27	534.4	16.2	644	6	CB862524	CB862524 HH04B08u
C 28	534	16.2	734	1	AL048427	AL048427 DKF2P566H
C 29	529	16.0	740	1	AL042909	AL042909 DKF2P434J
C 30	528	16.0	720	6	CA881974	CA881974 K0994C10-
C 31	523.2	15.9	528	1	AL038811	AL038811 DKF2P566O
C 32	519.4	15.7	521	6	CA890032	CA890032 B0156F05-
C 33	519	15.7	607	6	CB862099	CB862099 HH06A02y
C 34	516.4	15.6	518	6	CA886082	CA886082 B0124F04-
C 35	510.2	15.5	686	1	AL044407	AL044407 DKF2P434D
C 36	509.4	15.4	511	6	CA895436	CA895436 B0192A07-
C 37	498	15.1	498	1	AL039076	AL039076 DKF2P566G
C 38	494.8	15.0	1025	1	AL038025	AL038025 DKF2P566C
C 39	487.6	14.8	498	7	CK394397	CK394397 hggada2H1
C 40	477.6	14.5	491	6	CB862715	CB862715 HH03F14u
C 41	465.8	14.1	679	1	AL039128	AL039128 DKF2P566K
C 42	462.8	14.0	490	1	AL039589	AL039589 DKF2P434D
C 43	458	13.9	458	6	CA887583	CA887583 B0137H04-
C 44	456.4	13.8	471	1	AL039649	AL039649 DKF2P434G
C 45	449.2	13.6	612	6	CB863409	CB863409 HH04K05w

ALIGNMENTS

RESULT 1
LOCUS BQ751936 794 bp mRNA linear EST 18-JUL-2002
DEFINITION EST632499 DSCT Colletotrichum trifolii cDNA clone pDSCT8-67, mRNA
sequence.
ACCESSION BQ751936
VERSION BQ751936.1 GI:21907341
KEYWORDS EST.
SOURCE Colletotrichum trifolii
ORGANISM Colletotrichum trifolii
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
mitosporic Phyllachoraceae; Colletotrichum.
REFERENCE 1 (bases 1 to 794)
AUTHORS Samad D.A., Dickman, M., Town, C.D., Van Aken, S., Utterback, T.,
Cheung, F. and Fraser, C.M.
ESTs from mycelia of Colletotrichum trifolii race 1
Other ESTs: EST632498
Unpublished (2002)
CONTACT: Deborah A. Samac
Department of Plant Pathology
University of Minnesota
495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
Tel: 612 625 1243
Fax: 651 649 5058
Email: debbys@puccini.crl.umn.edu
TIGR sequence name: MTSAH67TV More information is available at:
www.medicago.org
Seq primer: (gtA AtA CgA Ctc Act AtA ggg C).
Location/Qualifiers
1..794
/organism="Colletotrichum trifolii"
/mol_type="mRNA"
/strain="race 1"
/db_xref="taxon:5466"
/clone="pDSCT8-67"
/tissue_type="mycelia"
/development="young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
/lab_host="DH5alpha"
/clone_lib="DSCT"

FEATURES
source

note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2sp2 ; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gtl1 from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce.

COMMENT

Other_ESTs: EST632217
 Contact: Deborah A. Samac
 Department of Plant Pathology
 University of Minnesota
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
 Tel: 612 625 1243
 Fax: 651 649 5058
 Email: debbys@puccini.crl.umn.edu
 TIGR sequence name: MTSAC64TV More information is available at:
 www.medicago.org
 Seq primer: (gta ata cga ctc act ata 999 C).

FEATURES

Location/Qualifiers
 1..793
 /organism="Colletotrichum trifolii"
 /mol_type="mRNA"
 /strain="race 1"
 /db_xref="taxon:5466"
 /clone="pDST6-59"
 /tissue_type="mycelia"
 /dev_stage="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
 /lab_host="DH5alpha"
 /clone_lib="DSTCT"
 /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2sp2; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gt11 from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN

Query Match 23.1%; Score 762; DB 5; Length 793;
 Best Local Similarity 99.1%; Pred. No. 3.3e-213;
 Matches 777; Conservative 0; Mismatches 5; Indels 2; Gaps 1;
 3 GACACATCGAATGGTGCAAAACCTTCGCGGTATGGCATATAGCCCGGAGAGAGT 62
 1 GACATCTTCGAATGCGGCAAAACCTTCGCGGTATGGCATATAGCCCGGAGAGAGT 60
 63 CAATTCAGGGTGGTGAATGTGAACACCAAGTAACTATACGATGTCGAGAGTATGCCGGT 122
 61 CAATTCAGGGTGGTGAATGTGAACACCAAGTAACTATACGATGTCGAGAGTATGCCGGT 120
 123 GTCTCTTATCAGACCGTTTCCGCGGTGGTGAACAGGCCAGCCACGTTTCTGCGAAAACG 182
 121 GTCTCTTATCAGACCGTTTCCGCGGTGGTGAACAGGCCAGCCACGTTTCTGCGAAAACG 180
 183 CGGGAAAAAGTGAAGCGGCATCGCGAGCTGAATTAACATTCCTCCACCGGTGGCAAA 242
 181 CGGGAAAAAGTGAAGCGGCATCGCGAGCTGAATTAACATTCCTCCACCGGTGGCAAA 240
 243 CAATCGCGGCAAAACAGTGTCTGATGGCGTTGCCACCTCCAGTCTGGCCCTGCAC 302
 241 CAATCGCGGCAAAACAGTGTCTGATGGCGTTGCCACCTCCAGTCTGGCCCTGCAC 300
 303 CGCGCGTGCAGAAATGTTCGGGGCATTAATCTCGCCCGCATCACTGGGTGCCAGGGT 362
 301 CGCGCGTGCAGAAATGTTCGGGGCATTAATCTCGCCCGCATCACTGGGTGCCAGGGT 360
 363 GTGGTGTGATGGTGAACAGCGGCTCGAAGCCTGTAAGCGGCGGTGCACATCTT 422
 361 GTGGTGTGATGGTGAACAGCGGCTCGAAGCCTGTAAGCGGCGGTGCACATCTT 420
 423 CTGCGCAACCGCTCACTGGGTGATCAATTAATCTCCGTGGATGACCAAGATGCCATT 482
 421 CTGCGCAACCGCTCACTGGGTGATCAATTAATCTCCGTGGATGACCAAGATGCCATT 480

QY 483 GCTGTGGAAGCTGCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTCCTTGACAGACA 542
 Db 481 GCTGTGGAAGCTGCTGCACTAAATGTTCCGGCGTTATTTCTTGATGTCCTTGACAGACA 540
 QY 543 CCCATCAACAGTATATTTCTCCCATGAAGACGGTACGGACTGGGGTGGAGCATCTG 602
 Db 541 CCCATCAACAGTATATTTCTCCCATGAAGACGGTACGGACTGGGGTGGAGCATCTG 600
 QY 603 GTCGCATTTGGTCAACCAAAATCGCGCTGTAGCGGCCCATTTAAGTCTTGTCTCGGCG 662
 Db 601 GTCGCATTTGGTCAACCAAAATCGCGCTGTAGCGGCCCATTTAAGTCTTGTCTCGGCG 660
 QY 663 CGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 722
 Db 661 CGTCTGCGTCTGGCTGGCATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG 720
 QY 723 GAACGGGAAGCGCACTGGAGTGCATGTCGGGTTTCAACAAACCATGCAAAATCTGAAT 782
 Db 721 GAACGGGAAGCGCACTGGAGTGCATGTCGGGTTTCAACAAACCATGCAAAATCTGAAT 778
 QY 783 GAGG 786
 Db 779 GAGG 782

RESULT 5
 BQ751220 806 bp mRNA linear EST 18-JUL-2002
 LOCUS EST631783 DSCT Colletotrichum trifolii cDNA clone pDST3-64, mRNA
 DEFINITION sequence.
 ACCESSION BQ751220
 VERSION BQ751220.1 GI:21906625
 KEYWORDS EST
 SOURCE Colletotrichum trifolii
 ORGANISM Colletotrichum trifolii
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 Sordariomycetes incertae sedis; Phyllachorales; Phyllachoraceae;
 Mitosporic Phyllachoraceae; Colletotrichum.
 REFERENCE 1 (bases 1 to 806)
 AUTHORS Samac,D.A., Dickman,M., Town,C.D., Van Aken,S., Utterback,T.,
 Cheung,F. and Frazer,C.M.
 TITLE ESTs from mycelia of Colletotrichum trifolii race 1
 JOURNAL Unpublished (2002)
 COMMENT Other ESTs: EST631782
 Contact: Deborah A. Samac
 Department of Plant Pathology
 University of Minnesota
 495 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108, USA
 Tel: 612 625 1243
 Fax: 651 649 5058
 Email: debbys@puccini.crl.umn.edu
 TIGR sequence name: MTSAC64TV More information is available at:
 www.medicago.org
 Seq primer: (gta ata cga ctc act ata 999 C).

FEATURES

Location/Qualifiers
 1..806
 /organism="Colletotrichum trifolii"
 /mol_type="mRNA"
 /strain="race 1"
 /db_xref="taxon:5466"
 /clone="pDST3-64"
 /tissue_type="mycelia"
 /dev_stage="Young, actively growing mycelia (3 days after inoculation) grown in liquid culture (cutin minimal medium containing 2% glucose)."
 /lab_host="DH5alpha"
 /clone_lib="DSTCT"
 /note="Vector: pBluescript SK+; Site 1: EcoRI; Site 2: EcoRI; isolate: 2sp2; cDNA was prepared from polyA+ enriched RNA The cDNA was ligated into Lambda gt11 from Stratagene and packaged using Gigapack packaging extracts. An aliquot of the amplified library was used to transduce E. coli Y1090 and phage DNA was purified from a liquid

lysate. The cDNA inserts were gel purified after EcoRI digestion and ligated into pBluescript SK+. Aliquots of the ligation were used to transform E. coli DH5alpha which were plated onto medium with X-gal for selection of recombinants."

ORIGIN		Query Match		Best Local Similarity		Matches		Conservative		0; Mismatches		16; Indels		5; Gaps		2;	
Qy	3	GACACATCGAATGTGTCGCAAAACCTTTTCGCGTATGTCATAGCGCCGGAAGAGAGT	62														
Db	1	GACACTTTCGAATGCGGCANAAACCTTTTCGCGTATGTCATAGCGCCGGAAGAGAGT	60														
Qy	63	CAATTCAGGGTGGTGAATGTGAAACACAGTAACTTATACGATGTCGCGAGATGCCGGT	122														
Db	61	CAATTCAGGGTGGTGAATGTGAAACACAGTAACTTATACGATGTCGCGAGATGCCGGT	120														
Qy	123	GTCTCTTATCAGACGTTTCCGCGTGTGAACAGGCGCAGGTCGTTTCTCGGAAAACG	182														
Db	121	GTCTCTTATCAGACGTTTCCGCGTGTGAACAGGCGCAGGTCGTTTCTCGGAAAACG	180														
Qy	183	CGGAAAAGTGAAGCGCGATGCGGAGCTGAATTAATTCCTCCAAACCGGTGGCAAA	242														
Db	181	CGGAAAAGTGAAGCGCGATGCGGAGCTGAATTAATTCCTCCAAACCGGTGGCAAA	240														
Qy	243	CAACTGGCGGCAAAACAGTCGTTGCTGATGTCGCGTGTGCACCTCCAGTCTGGCCCTGCAC	302														
Db	241	CAACTGGCGGCAAAACAGTCGTTGCTGATGTCGCGTGTGCACCTCCAGTCTGGCCCTGCAC	300														
Qy	303	GGCGCGTCGCAAAATGTCGCGCGATTAATCTCCGCGCGATCAACTGGGTGCCAGCGTG	362														
Db	301	GGCGCGTCGCAAAATGTCGCGCGATTAATCTCCGCGCGATCAACTGGGTGCCAGCGTG	360														
Qy	363	GTGGTGTGATGATAGAACGCGCTGCGAGCGCTGAAGCGCGGTGCACAACTCTT	422														
Db	361	GTGGTGTGATGATAGAACGCGCTGCGAGCGCTGAAGCGCGGTGCACAACTCTT	420														
Qy	423	CTCGCGCAACGCGTCAGTGGCTGATCAATTAATCTCGCTGGATGACCGAGATGCCATT	482														
Db	421	CTCGCGCAACGCGTCAGTGGCTGATCAATTAATCTCGCTGGATGACCGAGATGCCATT	480														
Qy	483	GCTGTGGAAGCTGCTGCACAAATGTCGCGGTTATTTCTTGATGTCTTGACGACACA	542														
Db	481	GCTGTGGAAGCTGCTGCACAAATGTCGCGGTTATTTCTTGATGTCTTGACGACACA	540														
Qy	543	CCCATCAACAGTATTTATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATCTG	602														
Db	541	CCCATCAACAGTATTTATTTCTCCCATGAAGCGGTACGCGACTGGGCGTGGAGCATCTG	600														
Qy	603	GTGCGATTGGGTCAACAGCAAAATCGGCTGTTAGCGGGCCCAATTAAGTTCTGTCGCGG	662														
Db	601	GTGCGATTGGGTCAACAGCAAAATCGGCTGTTAGCGGGCCCAATTAAGTTCTGTCGCGG	660														
Qy	663	CGTCTGCGTCTGGCTGGCGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG	722														
Db	661	CGTCTGCGTCTGGCTGGCGATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCG	720														
Qy	723	GAACGGGAAGGCGACTGGAGTCCCATGTCGCGTTTTCACAAACCATGCAATGCTGAAT	782														
Db	721	GAACGGGAAGGCGACTGGAGTCCCATGTCGCGTTTTCACAAACCATGCAATGCTGAAT	777														
Qy	783	GAGGCGATCGTTCCCACTGCGATGCTGGTTG	813														
Db	778	GAGG---CATCGTTCACCTCGATGCTGGTTG	806														

RESULT 6
CB863541/c
LOCUS
DEFINITION HH04A08y HH Hordeum vulgare cDNA clone HH04A08 3-PRIME, mRNA
689 bp mRNA linear EST 22-APR-2003
sequence.

ACCESSION	CB863541	GI:30058100
VERSION	EST.	
KEYWORDS	Hordeum vulgare	
SOURCE	Hordeum vulgare	
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae; Triticeae; Hordeum.	
REFERENCE	1. (bases 1 to 689)	
AUTHORS	Varshney,R.K., Zhang,H., Burton,R., Stein,N., Langridge,P. and Graner,A.	
TITLE	Barley ESTs from coleoptile tissue	
JOURNAL	Unpublished (2003)	
COMMENT	Contact: Stein Mills Molecular Markers Group, Department Genbank Institute of Plant Genetics and Crop Plant Research (IPK) Corrensstr. 3, 06466, Gatersleben, Germany Tel: 039482-5522 Fax: 039482-5595 Email: stein@ipk-gatersleben.de Insert Length: 689 Std Error: 0.00 Plate: 4 row: A column: 8 Seq primer: SP6.	
FEATURES	Location/Qualifiers	
source	1..689	
	/organism="Hordeum vulgare"	
	/mol_type="mRNA"	
	/cultivar="Sloop"	
	/db_xref="GABI:555198"	
	/db_xref="taxon:4513"	
	/clone="HH04A08"	
	/tissue_type="coleoptile"	
	/dev_stage="coleoptile, 1 day old"	
	/lab_host="DH10B"	
	/clone_lib="HH"	
	/note="Vector: pSPORT; Site 1: SalI (5-end of cDNA); Site 2: NotI (3-end of cDNA); Due to the cloning system used blue/white selection for recombinants is not 100 % reliable. Average insert size is 1.3 kb."	
ORIGIN	Query Match	20.8%; Score 684.8; DB 6; Length 689;
	Best Local Similarity	99.6%; Pred. No. 2.1e-190;
	Matches	686; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy	21	AAAACCTTTTCGCGTATGCGCATGATAGCGCCGGAAGAGAGTCAATTCAGGTGGTGAAT 80
Db	689	AAAACCTTTTCGCGTATGCGCATGATAGCGCCGGAAGAGAGTCAATTCAGGTGGTGAAT 630
Qy	81	GTGAAACCAAGTAACGTTATACGATGTCGAGAGTATGCGGTCTCTTATCAGACCGTT 140
Db	629	GTGAAACCAAGTAACGTTATACGATGTCGAGAGTATGCGGTCTCTTATCAGACCGTT 570
Qy	141	TCCCGCGTGTGAAACCAAGTATGCGCATGATAGCGCCGGAAGAGAGTCAATTCAGGTGGTGAAT 200
Db	569	TCCCGCGTGTGAAACCAAGTATGCGCATGATAGCGCCGGAAGAGAGTCAATTCAGGTGGTGAAT 510
Qy	201	GCGATGGCGAGTCAATTAATTAATTCATTCCTCCAAACCGCGTGGCAACAACCTGCGCGGCAACAG 260
Db	509	GCGATGGCGAGTCAATTAATTAATTCATTCCTCCAAACCGCGTGGCAACAACCTGCGCGGCAACAG 450
Qy	261	TGTTGCTGATGGCGTTGCCACCTTCCAGTCTGGCCCTTCAGCGCCGCGTTCGCAAAATTGTC 320
Db	449	TGTTGCTGATGGCGTTGCCACCTTCCAGTCTGGCCCTTCAGCGCCGCGTTCGCAAAATTGTC 390
Qy	321	GCGCGATTAAATCTCGCGCGGATCAACTGGGTGCCAGGTGGTGGTTCGATGCTAGTA 380
Db	389	GCGCGATTAAATCTCGCGCGGATCAACTGGGTGCCAGGTGGTGGTTCGATGCTAGTA 330
Qy	381	CGAAGCGCGTGAAGAGCTGTAAAGCGCGGTGCAACAATCTTCTCGCGCAACGCGTCACT 440
Db	329	CGAAGCGCGTGAAGAGCTGTAAAGCGCGGTGCAACAATCTTCTCGCGCAACGCGTCACT 270

/mol_type="genomic DNA"
/db_xref="taxon:186617"
/clone="SI051p3G6L"
/clone_lib="Scripps Pier (La Jolla, CA) uncultured virus community"
/note="Marine viruses were isolated from 200 liters of surface seawater using a combination of differential filtration and density-dependent gradient centrifugation. Linker-amplified shotgun libraries were constructed by randomly shearing the total marine viral community DNA, end-repairing, ligating dsDNA linkers to the ends, and amplifying the fragments using Vent DNA polymerase. The resulting fragments were ligated into the pSMART vector and electroporated into MC12 cells (Lucigen; Middleton, WI)"

ORIGIN

Query Match 20.5%; Score 676; DB 8; Length 710;
Best Local Similarity 98.9%; Pred. No. 8.4e-188;
Matches 590; Conservative 0; Mismatches 7; Indels 1; Gaps 1;

Qy 400 GTAAAGCGCGGTGCACAAATCTTCTCGCGCAACCGCTCAGTGGGTGATCAATTAACATATC 459
Db |||
8 GTGCTGCAGCGGTGCACAAATCTTCTCGCGCAACCGCTCAGTGGGTGATCAATTAACATATC 67
Qy 460 CGCTGGATGACCGATGCCATTGCTGTGGAAGCTGCTGCACATATGTTCCGGCGTTAT 519
Db |||
68 CGCTGGATGACCGATGCCATTGCTGTGGAAGCTGCTGCACATATGTTCCGGCGTTAT 127
Qy 520 TTCTTGATGTCTCTGACGACACCCATCAACAGTATTATTTCTCCCATGAAGACGGTA 579
Db |||
128 TTCTTGATGTCTCTGACGACACCCATCAACAGTATTATTTCTCCCATGAAGACGGTA 187
Qy 580 CGCGACTGGCGGTGAGCATCTGCTGCGATTTGGGTCAACAGAAATCGCGTGTAGCGG 639
Db |||
188 CGCGACTGGCGGTGAGCATCTGCTGCGATTTGGGTCAACAGAAATCGCGTGTAGCGG 247
Qy 640 GCCCATTAAGTCTGTCTCGGGCGTCTGCTGCTGGCTGGCTGGCATAAATATCTCACTC 699
Db |||
248 GCCCATTAAGTCTGTCTCGGGCGTCTGCTGCTGGCTGGCTGGCATAAATATCTCACTC 307
Qy 700 GCAATCAAAATTCAGCGGATAGCGGAACGGAAGGCGACTGGAGTGCCTATGTCGGGTTTC 759
Db |||
308 GCAATCAAAATTCAGCGGATAGCGGAACGGAAGGCGACTGGAGTGCCTATGTCGGGTTTC 367
Qy 760 AACAAACCATGCAATGCTGAATGAGGCGATCGTTCCCACTGCGATGCTGTTGCCAACG 819
Db |||
368 AACAAACCATGCAATGCTGAATGAGGCGATCGTTCCCACTGCGATGCTGTTGCCAACG 427
Qy 820 ATCAGATGGCGTGGCGCAATGCGCGCAATTACCGAGTCCGGGTGCGGTGTTGTCGG 879
Db |||
428 ATCAGATGGCGTGGCGCAATGCGCGCAATTACCGAGTCCGGGTGCGGTGTTGTCGG 487
Qy 880 ATATCTCGGTAGTGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAA 939
Db |||
488 ATATCTCGGTAGTGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCGTTAA 547
Qy 940 CCACCATCAACAGGATTTTCGCTGCTGGGCAACACGAGTGGACGCTTGTGTCAC 999
Db |||
548 CCACCATCAACAGGATTTTCGCTGCTGGGCAACACGAGTGGACGCTTGTGTCAC 607
Qy 1000 TCTCTCAGGGCGAGCGGTGAAGGCAATCAGCTGTTGCGGTCTCACTGTTGAAAGAA 1059
Db |||
608 TCTCTCAGGGCGAGCGGTGAAGGCAATCAGCTGTTGCGGTCTCACTGTTGAAAGAA-AA 666
Qy 1060 AAACCAACCTGGCGCCCAATAGCAAAACCGCTCTCCC 1097
Db |||
667 AAACCAACCTGGCGCCCAATAGCAAAACCGNCTTTCCC 704

RESULT 9
AL044483/c
LOCUS AL044483 660 bp mRNA linear EST 04-SEP-2003

DEFINITION DKFZp4341102.s1.434 (synonym: htes3) Homo sapiens cDNA clone
DKFZp4341102.3, mRNA sequence.
ACCESSION AL044483
VERSION AL044483.1 GI:5432701
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 660)
AUTHORS Ansorge, W., Benes, V., Krieger, S., Mewes, H.W., Gassenhuber, J. and Wismann, S.
TITLE EST (Ansorge, Benes, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 3' sequence of the clone insert
Clone from S. Wismann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wismann@dkfz-heidelberg.de;
sequenced by EMBL (European Molecular Biology Laboratories,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No r1 sequence available.
This clone (DKFZp4341102) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..660
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp4341102"
/tissue type="testis"
/dev stage="adult"
/lab host="DH10B"
/clone lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 19.6%; Score 647.4; DB 1; Length 660;
Best Local Similarity 99.7%; Pred. No. 2.4e-179;
Matches 659; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Qy 85 AACAGTAACGTTATACGATGTCGACAGTATGCGGTGTTCTTTATCAGACCGTTCCC 144
Db |||
660 AACAGTAACGTTATACGATGTCGACAGTATGCGGTGTTCTTTATCAGACCGTTCCC 601
Qy 145 CGGTGGTGAACAGCCAGCCACGTTTCTGCGAAAACGCGGGAAGTGGAGCGGCA 204
Db |||
600 CGGTGGTGAACAGCCAGCCAGCCACGTTTCTGCGAAAACGCGGGAAGTGGAGCGGCA 541
Qy 205 TGGCGGAGCTCAATTACATTCCTCAACCGCTGGGACACAACTGCGGGGCAACAGTCGT 264
Db |||
540 TGGCGGAGCTCAATTACATTCCTCAACCGCTGGGACACAACTGCGGGGCAACAGTCGT 481
Qy 265 TGCTGATTTGGGTTGCCACCTCCAGTCTGGCCCTTCACCGCCCGTTCGAAATTTGTCGGG 324
Db |||
480 TGCTGATTTGGGTTGCCACCTCCAGTCTGGCCCTTCACCGCCCGTTCGCAAAATTTGTCGGG 421
Qy 325 CGATTAAATCTCGCGCCGATCAACTGGGTGCGAGGTGGTGTGATGTTAGAACGAA 384
Db |||
420 CGATTAAATCTCGCGCCGATCAACTGGGTGCGAGGTGGTGTGATGTTAGAACGAA 361
Qy 385 CGCGCTCGAAGCTGTAAAGCGCGGTGCACAACTTCTTCGCGCAACGCGTCAGTGGGC 444
Db |||
360 CGCGCTCGAAGCTGTAAAGCGCGGTGCACAACTTCTTCGCGCAACGCGTCAGTGGGC 301
Qy 445 TGATCAATTAATCTCCGCTGGATGACCAAGATGTCATTGCTGTGGAAGCTGCCTGCACTA 504
Db |||
300 TGATCAATTAATCTCCGCTGGATGACCAAGATGTCATTGCTGTGGAAGCTGCCTGCACTA 241
Qy 505 ATGTTTCGGCGGTTATTTCTTGTGATGTCTCTGACGACACCCATCAACAGTATTTTCT 564

```

Db      240  ATGTTCCGGCGTTATTCTTGTGATGCTCTGACACAGACCCCATCAACAGTATTATTCT 181
QY      565  CCATGAAGACGGTAGCGACTGGCGGTGGAGCATCTGTCGATTTGGGTACACGACAAA 624
Db      180  CCATGAAGACGGTAGCGACTGGCGGTGGAGCATCTGTCGATTTGGGTACACGACAAA 121
QY      625  TCGCGCTGTTAGCGGCCCAATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGC 684
Db      120  TCGCGCTGTTAGCGGCCCAATTAAGTTCTGTCTCGGCGCGTCTGCGTCTGGCTGGCTGGC 62
QY      685  ATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGCGGACTGGAGTG 744
Db      61  ATAAATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGCGGACTGGAGTG 2
QY      745  C 745
Db      1  C 1

RESULT 10
BJ063872/c
LOCUS
DEFINITION
BJ063872 NIBB Mochii normalized Xenopus tailbud library Xenopus
laevis cDNA clone XL07K02 5', mRNA sequence.
ACCESSION
BJ063872
VERSION
BJ063872.1 GI:17471074
KEYWORDS
EST.
SOURCE
Xenopus laevis (African clawed frog)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Xenopus.
1 (bases 1 to 641)
REFERENCE
Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i.T. and
Kohara,Y.
Expressed genes in X. laevis embryo
Unpublished (2001)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp
The information of this clone is available through the following
URL.
http://xenopus.nibb.ac.jp.
FEATURES
Location/Qualifiers
1..641
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="XL07K02"
/tissue_type="whole embryo"
/dev stage="stage 25"
/clone_lib="NIBB Mochii normalized Xenopus tailbud
library"
ORIGIN
Query Match 19.4%; Score 641; DB 4; Length 641;
Best Local Similarity 100.0%; Pred. No. 1.8e-177;
Matches 641; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      30  CGCGGTATGCATGATAGCCCGGAGAGAGATCAATTCAGGTGTTGATGTGAACCA 89
Db      641  CGCGGTATGCATGATAGCCCGGAGAGAGATCAATTCAGGTGTTGATGTGAACCA 582
QY      90  GTAAAGTTATACGATGTCGACAGATAGCCGGTGTCTTTATCAGACCGTTTCCCGGTG 149
Db      581  GTAAAGTTATACGATGTCGACAGATAGCCGGTGTCTTTATCAGACCGTTTCCCGGTG 522
QY      150  GTGAACGACGCCAGCCAGTTTCTGCGAAGAACCGGGGAAAAAGTGAAGCGCGATGGCG 209

```

```

Db      521  GTGAACGACGCCAGCCAGTTTCTCGAAGAACCGGGGAAAAAGTGAAGCGCGATGGCG 462
QY      210  GAGCTGAATTACATTCCCAACCGGTGGCAACAACCTGGGGGCAACAGTTCGTTGCTG 269
Db      461  GAGCTGAATTACATTCCCAACCGGTGGCAACAACCTGGGGGCAACAGTTCGTTGCTG 402
QY      270  ATTGGCGTTGCACTCCAGTCTGGCCCTGACGCGCGCTGCGCAAAATTGTGCGCGCGATT 329
Db      401  ATTGGCGTTGCACTCCAGTCTGGCCCTGACGCGCGCTGCGCAAAATTGTGCGCGCGATT 342
QY      330  AAATCTCGCGCGCATCAACTGGGTGCCAGCGTGGTGTGTCGATGTTAGAACGAGCGCG 389
Db      341  AAATCTCGCGCGCATCAACTGGGTGCCAGCGTGGTGTGTCGATGTTAGAACGAGCGCG 282
QY      390  GTCGAAGCCTGTAAAGCGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATC 449
Db      281  GTCGAAGCCTGTAAAGCGCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATC 222
QY      450  ATTAACCTATCGCTGGATGACAGGATGCCATTTGTTGGAAGCTGCTGCACCTAATGTT 509
Db      221  ATTAACCTATCGCTGGATGACAGGATGCCATTTGTTGGAAGCTGCTGCACCTAATGTT 162
QY      510  CCGCGCTTATTTCTTGATGTTCTTGACCAACACCCATCAACAGTATTATTTTCTCCCAT 569
Db      161  CCGCGCTTATTTCTTGATGTTCTTGACCAACACCCATCAACAGTATTATTTTCTCCCAT 102
QY      570  GAAGACGGTACGGCACTGGGCGTGGAGCATCTGTCGCAATGGGTACCAGCAAAATCGCG 629
Db      101  GAAGACGGTACGGCACTGGGCGTGGAGCATCTGTCGCAATGGGTACCAGCAAAATCGCG 42
QY      630  CTGTTAGCGGCGCCATTAAAGTTCTGTTCTGGCGCGTCTGGG 670
Db      41  CTGTTAGCGGCGCCATTAAAGTTCTGTTCTGGCGCGTCTGGG 1

RESULT 11
CB863814/c
LOCUS
DEFINITION
HH07B20y HH Hordeum vulgare cDNA clone HH07B20 3-PRIME, mRNA
sequence.
ACCESSION
CB863814
VERSION
CB863814.1 GI:30058373
KEYWORDS
EST.
SOURCE
Hordeum vulgare
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
1 (bases 1 to 624)
Varshney,R.K., Zhang,H., Burton,R., Stein,N., Langridge,P. and
Graner,A.
Barley ESTs from coleoptile tissue
Unpublished (2003)
Contact: Stein Nils
Molecular Markers Group, Department Genbank
Institute of Plant Genetics and Crop Plant Research (IPK)
Corrensstr. 3, 06466, Gatersleben, Germany
Tel: 039482-5522
Fax: 039482-5595
Email: stein@ipk-gatersleben.de
Insert Length: 624 Std Error: 0.00
Plate: 7 row: B column: 20
Seq primer: SP6.
Location/Qualifiers
1..624
/organism="Hordeum vulgare"
/mol_type="mRNA"
/cultivar="Sloop"
/db_xref="GABI:555471"
/db_xref="taxon:4513"
/clone="HH07B20"
/tissue_type="coleoptile"
FEATURES
source

```

/dev_stage="coleoptile, 1 day old"
/lab_host="DH10B"
/clone_lib="HH"
/note="Vector: pSPORT; Site 1: SalI (5-end of cDNA);
Site 2: NotI (3-end of cDNA); Due to the cloning system
used 'blue/white selection for recombinants is not 100 %
reliable. Average insert size is 1.3 Kb."

ORIGIN

Query Match 18.7%; Score 615.8; DB 6; Length 624;
Best Local Similarity 99.7%; Pred. No. 5.1e-170;
Matches 617; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 GACACATCGAATGTGCAAAACCTTTTCGCGTATGGCATGTAGCGCCGCGAAGAGAGT 62
DB 619 GACACATCGAATGTGCGCAAAACCTTTTCGCGTATGGCATGTAGCGCCGCGAAGAGAGT 560

QY 63 CAATTCAGGGTGTCAATGTGAAACAGTAACGTTATACGATGTCGACAGATATCCCGGT 122
DB 559 CAATTCAGGGTGTGAATGTGAAACAGTAACGTTATACGATGTCGACAGATATCCCGGT 500

QY 123 GTCTCTTATCAGACCGTTTCCCGCGTGTGTAACACGAGCCAGCGTTCGCGAAAAACG 182
DB 499 GTCTCTTATCAGACCGTTTCCCGCGTGTGTAACACGAGCCAGCGTTCGCGAAAAACG 440

QY 183 CGGAAAAAGTGAAGCGCGGATGCGGAGCTGAATTACATTTCCCAACCGCGTGCGACAA 242
DB 439 CGGAAAAAGTGAAGCGCGGATGCGGAGCTGAATTACATTTCCCAACCGCGTGCGACAA 380

QY 243 CAACCTGGCGGCAACAGCTGCTGATTTGGCGTTGCCACCTCCAGTCTGCCCTGCAC 302
DB 379 CAACCTGGCGGCAACAGCTGCTGATTTGGCGTTGCCACCTCCAGTCTGCCCTGCAC 320

QY 303 GCGCGTGTGCAAAATGTGCGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCGTG 362
DB 319 GCGCGTGTGCAAAATGTGCGCGGATTAATCTCGCGCGATCAACTGGGTGCCAGCGTG 260

QY 363 GTGGTGTGATGTTAGAACGAGCGCGTTCGAAGCCTGTAAAGCGCGGTGCACAACTTT 422
DB 259 GTGGTGTGATGTTAGAACGAGCGCGTTCGAAGCCTGTAAAGCGCGGTGCACAACTTT 200

QY 423 CTCGCGCAACGGTCAGTGGGCTGATCAATTAATCTCGCGGATCACTCCGCTGGATGCCAT 482
DB 199 CTCGCGCAACGGTCAGTGGGCTGATCAATTAATCTCGCGGATCACTCCGCTGGATGCCAT 140

QY 483 GCTGTGGAAGCTGCTGCACATAATGTTCCGGGTTATTTCTTGTGATGTTCTCTCACAGACA 542
DB 139 GCTGTGGAAGCTGCTGCACATAATGTTCCGGGTTATTTCTTGTGATGTTCTCTCACAGACA 80

QY 543 CCCATCAACAGTATTTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATCTG 602
DB 79 CCCATCAACAGTATTTTCTCCATGAAGACGGTACGCGACTGGCGGTGGAGCATCTG 20

QY 603 GTCGATTTGGGTACACAGC 621
DB 19 GTCGATTTGGGTACACAGC 1

RESULT 12
AL045353/c
LOCUS
DEFINITION DKF2p34B075.s1.434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION AL045353
VERSION AL045353.1 GI:5433506
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 794)
Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.

TITLE EST (Duesterhoeft, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS

MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 3' sequence of the clone insert.
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); sequenced by Qiagen within the cDNA
sequencing consortium of the German Genome Project. r1 sequence
also available.

This clone is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

source
1..794
location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKF2p34B075"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI"

ORIGIN

Query Match 18.5%; Score 612; DB 1; Length 794;
Best Local Similarity 99.7%; Pred. No. 7.3e-169;
Matches 634; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 569 TGAAGACGGTACGACCTGGCGGTGGAGCATCTGTCGCATTTGGTCAACGAAATCGC 628
DB 794 TGAAGACGGTACGACCTGGCGGTGGAG-ATCTGGTCGATTTGGTCAACGAAATCGC 736

QY 629 GCTGTTAGCGGCGCATTAAGTCTGCTCGCGCGTCTGCTGGTCTGGTGGTGCATAA 688
DB 735 GCTGTTAGCGGCGCATTAAGTCTGCTCGCGCGTCTG-GTCTGGTGGTGGCATAA 677

QY 689 ATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGCGCATTTGAGTGCAT 748
DB 676 ATATCTCACTCGCAATCAAAATTCAGCCGATAGCGGAACGGGAAGCGCATTTGAGTGCAT 617

QY 749 GTCGCGTTTTCAACAAACCATGCAAAATGCTGAATGAGGGCATTCGTTCCCATCGCATGCT 808
DB 616 GTCGCGTTTTCAACAAACCATGCAAAATGCTGAATGAGGGCATTCGTTCCCATCGCATGCT 557

QY 809 GGTTCGCCAACCATCAGATGGCGCTGGCGGCAATCGCGCCATTTACCGAGTCCGGCTGCG 868
DB 556 GGTTCGCCAACCATCAGATGGCGCTGGCGGCAATCGCGCCATTTACCGAGTCCGGCTGCG 497

QY 869 GCTTGGTCGGATATCTCGGTAGTGGGATAGCAGATACCGAAGACAGCTCATGTTATAT 928
DB 496 GCTTGGTCGGATATCTCGGTAGTGGGATAGCAGATACCGAAGACAGCTCATGTTATAT 437

QY 929 CCCGCGTTAAACCATCAAAACAGGATTTTCGCTGCTGGGGCAAAACAGCGTGGACCG 988
DB 436 CCCGCGTTAAACCATCAAAACAGGATTTTCGCTGCTGGGGCAAAACAGCGTGGACCG 377

QY 989 CTTCGTGCAACTCTCTCAGGCGCAGCGGTGAGGGCAATCAGCTGTTGCCGCTCTCAT 1048
DB 376 CTTCGTGCAACTCTCTCAGGCGCAGCGGTGAGGGCAATCAGCTGTTGCCGCTCTCAT 317

QY 1049 GGTGAAAAAGAAAAACCCCTTGGGCGCCCAATACGCAAAACCGCTCTCCCGCGCGTTGGC 1108
DB 316 GGTGAAAAAGAAAAACCCCTTGGGCGCCCAATACGCAAAACCGCTCTCCCGCGCGTTGGC 257

QY 1109 CGATTCAATTAATGACGTGGCACGACAGTTTCCCGACTGGAAACCGGGCAGTAGCGCA 1168
DB 256 CGATTCAATTAATGACGTGGCACGACAGTTTCCCGACTGGAAACCGGGCAGTAGCGCA 197

QY 1169 ACGCAATTAATGTGAGTTAGTCTACTCATATTAGGCAC 1204
DB 196 ACGCAATTAATGTGAGTTAGTCTACTCATATTAGGCAC 161


```

RESULT 13
LOCUS CL655575/c
DEFINITION PRI0124a.C01 - PRI0124a.B21 (752) Mixed stage fosmid library of P.
pacificus var. California Pristionchus pacificus genomic, genomic
survey sequence.
ACCESSION CL655575
VERSION CL655575.1 GI:50134897
SOURCE GSS.
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida;
Neodiplogasteridae; Pristionchus.
REFERENCE 1 (bases 1 to 752)
AUTHORS Srinivasan,J., Otto G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
TITLE AppADB: an AcedB database for the nematode satellite organism
JOURNAL Pristionchus pacificus
COMMENT Nucleic Acids Res. 32 (1), D421-D422 (2004)
Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: raif.sommer@tuebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end
sequenced at Vancouver, Canada.
Seq primer: T7
Class: fosmid ends.

FEATURES             Location/Qualifiers
     source            1..752
     /organism="Pristionchus pacificus"
     /mol_type="genomic DNA"
     /strain="California"
     /db_xref="taxon:54126"
     /clone_lib="Mixed stage fosmid library of P. pacificus
var. California"
     /note="Vector: pEpifos-5 Fosmid vector"

ORIGIN
Query Match      18.3%; Score 603; DB 9; Length 752;
Best Local Similarity 98.9%; Pred. No. 3.3e-166;
Matches 628; Conservative 0; Mismatches 5; Indels 2; Gaps 2;

QY 1996 TTCACCTGGCGCTGATTGCTGCTGAC -GGGGGTTATGCGTTCAAGTATGAAACGGCAA 2054
Db 1996
QY 752 TTCACCTGGCGCTGATTGCTGCTGACGGGGGTTATGCGTTCAAGTATGAAACGGCAA 693
Db 752
QY 2055 GTACGACATTAAGACGTGGCGGTGGATTAACGCTGGCGCGAAACGGGTCTGACCTTCCT 2114
Db 2055
QY 692 GTACGACATTAAGACGTGGCGGTGGATTAACGCTGGCGCGAAACGGGTCTGACCTTCCT 633
Db 692
QY 2115 GGTTCACCTGATTAAAAACAACATGAATGACAGACCGATTACTCCATCGCAGAAC 2174
Db 2115
QY 632 GGTTCACCTGATTAAAAACAACATGAATGACAGACCGATTACTCCATCGCAGAAC 573
Db 632
QY 2175 TGCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCGCGTGGCGATGGTCCACAT 2234
Db 2175
QY 572 TGCCTTTAATAAGCGGAAACAGCGATGACCATCAACGGCGCGTGGCGATGGTCCACAT 513
Db 572
QY 2235 CGACACAGCAAGTGAATTATGTGTAAACGGTACTGCGACCTTCAAGGGTCAACCATC 2294
Db 2235
QY 512 CGACACAGCAAGTGAATTATGTGTAAACGGTACTGCGACCTTCAAGGGTCAACCATC 453
Db 512
QY 2295 CAACACCGTTCGTTGGCGTGTGACGCGAGGTATTAACCGCGCAGTCCGAAACAGAGCT 2354
Db 2295
QY 452 CAACACCGTTCGTTGGCGTGTGACGCGAGGTATTAACCGCGCAGTCCGAAACAGAGCT 393
Db 452
QY 2355 GGCAGAAAGAGTTCTCGAAACACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAA -TA 2413
Db 2355
QY 392 GGCAGAAAGAGTTCTCGAAACACTATCTGCTGACTGATGAAGGTCTGGAAGCGGTTAATTA 333

```

```

QY 2414 AAGACAAACCGCTGGTGGCCGTAGCTGAAAGTCTTACGAGGAAGAGTTGCGGAAAGATC 2473
Db 332 AGGCCAACCCGTTGGTGGCGCTAGCGTGAAGTCTTACGAGGAAGAGTTGCGGAAAGATC 273
QY 2474 CACGTATTGGCGCCACCATGAAACGCCCAAGAGTGAATCATGCCGAACATCCCGC 2533
Db 272 CACGTATTGGCGCCACCATGAAACGCCCAAGAGTGAATCATGCCGAACATCCCGC 213
QY 2534 AGATGTCCGCTTTCTGGTATGCCGTGCTACTGCGGTGATCAACGCCCGCCAGCGGTGTC 2593
Db 212 AGATGTCCGCTTTCTGGTATGCCGTGCTACTGCGGTGATCAACGCCCGCCAGCGGTGTC 153
QY 2594 AGATGTGTCATGAAGCCCTGAAGACGCGCAGACT 2628
Db 152 AGACTGTGTCATGAAGCCCTGAAGACGCGCAGACT 118

RESULT 14
LOCUS AY080094/c
DEFINITION AY080094 Scripps Pier (La Jolla, CA) uncultured virus community
uncultured marine virus genomic clone SIO51p3D4L, genomic survey
sequence.
ACCESSION AY080094
VERSION AY080094.1 GI:24745276
KEYWORDS GSS.
SOURCE uncultured marine virus
ORGANISM uncultured marine virus
REFERENCE 1 (bases 1 to 608)
AUTHORS Breitbart,M., Salamon,P., Andresen,B., Mahaffy,J.M., Segall,A.M.,
Mead,D., Azam,F. and Rohwer,F.
TITLE Genomic analysis of uncultured marine viral communities
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (22), 14250-14255 (2002)
MEDLINE 22294988
PUBMED 12384570
COMMENT Contact: Rohwer F
Biology Dept.
San Diego State University
5500 Campanile Dr, San Diego, CA 92102, USA
Tel: 6195941336
Fax: 619595676
Email: forest@unstroke.sdsu.edu
Class: shotgun.

FEATURES             Location/Qualifiers
     source            1..608
     /organism="uncultured marine virus"
     /mol_type="genomic DNA"
     /db_xref="taxon:186617"
     /clone="SIO51p3D4L"
     /clone_lib="Scripps Pier (La Jolla, CA) uncultured virus
community"
     /note="Marine viruses were isolated from 200 liters of
surface seawater using a combination of differential
filtration and density-dependent gradient centrifugation.
Linker-amplified shotgun libraries were constructed by
randomly shearing the total marine viral community DNA,
end-repairing, ligating dsDNA linkers to the ends, and
amplifying the fragments using Vent DNA polymerase. The
resulting fragments were ligated into the pSMART vector
and electroporated into MC12 cells (Lucigen; Middleton,
WI)"

ORIGIN
Query Match      17.9%; Score 589.8; DB 8; Length 608;
Best Local Similarity 99.7%; Pred. No. 2.5e-162;
Matches 591; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 287 CAGTCTGGCCCTGCACGCGCCCGTCGCAAAATTCGCGCGGATTAAATCTCGCGCGATCA 346
Db 608 CAGTCTGGCCCTGCACGCGCCCGTCGCAAAATTCGCGCGGATTAAATCTCGCGCGATCA 549

```

```
QY 347 ACTGGGTGCCAGCGTGGTGTCTCGATGTTAGAACGAAGCGCGTCGAAGCCTGTAAAGC 406
|
Db 548 ACTGGGTGCCAGCGTGGTGTCTCGATGTTAGAACGAAGCGCGTCGAAGCCTGTAAAGC 489
|
QY 407 GCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCAATACTATCCGCTGGA 466
|
Db 488 GCGGTGCACAATCTTCTCGCGCAACGCGTCAGTGGGCTGATCAATACTATCCGCTGGA 429
|
QY 467 TGACCAAGATGCCATGCTGTGGGAAGCTGCCGCACTAATGTTCCGGGCTTATTTCTTGA 526
|
Db 428 TGACCAAGATGCCATGCTGTGGGAAGCTGCCGCACTAATGTTCCAGGCTTATTTCTTGA 369
|
QY 527 TGTCTCTGACCAAGACACCCATCAACAGATATATTTCTCCCATGAAGACGCTACCGGACT 586
|
Db 368 TGTCTCTGACCAAGACACCCATCAACAGATATATTTCTCCCATGAAGACGCTACCGGACT 309
|
QY 587 GGGCGTGGAGCATCTGGTCGCAATTGGGTCAACAGCAAAATCGGCTGTTAGCGGGCCCAT 646
|
Db 308 GGGCGTGGAGCATCTGGTCGCAATTGGGTCAACAGCAAAATCGGCTGTTAGCGGGCCCAT 249
|
QY 647 AAGTTCTGTCTCGGCGGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCA 706
|
Db 248 AAGTTCTGTCTCGGCGGTCTGCGTCTGGCTGGCTGGCATAAATATCTCACTCGCAATCA 189
|
QY 707 AATTCAAGCCGATAGCGGAACGGAAGCGACTGGAGTGCCATGTCGGGTTTTCACAAAC 766
|
Db 188 AATTCAAGCCGATAGCGGAACGGAAGCGACTGGAGTGCCATGTCGGGTTTTCACAAAC 129
|
QY 767 CATGCAAAATGCTGAATGAGGCGCATGTTCCCACTGCGATGCTGTTGCCAAGATCAGAT 826
|
Db 128 CATGCAAAATGCTGAATGAGGCGCATGTTCCCACTGCGATGCTGTTGCCAAGATCAGAT 69
|
QY 827 GCGCTGGGCGCAATGCGGCGCATTTACCGAGTCCGGGCTGCGGTTGTTGCGG 879
|
Db 68 GCGCTGGGCGCAATGCGGCGCATTTACCGAGTCCGGGCTGCGGTTGTTGCGG 16
|
```

```
RESULT 15
AL037742/c
LOCUS
DEFINITION
  DKFZp564C157.r1 564 (synonym: hfbr2) Homo sapiens cDNA clone
  DKFZp564C157.5', mRNA sequence.
ACCESSION
  AL037742
VERSION
  AL037742.1 GI:5407088
KEYWORDS
  EST.
SOURCE
  Homo sapiens (human)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 589)
  Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
  Wiemann,S.
  EST (Bloecker, et al.)
  Unpublished (1999)
  Contact: MIPS
  MIPS
```

```
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by GBF (National Research Centre for Biotechnology Ltd.,
Braunschweig/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No s1 sequence available.
This clone (DKFZp564C157) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
  1..589
  /organism="Homo sapiens"
  /mol_type="mRNA"
  /db_xref="taxon:9606"
  /clone="DKFZp564C157"
```

FEATURES
source

```
/tissue_type="brain"
/dev_stages="fetal"
/lab_host="xl-2blue"
/clone_lib="564 (synonym: hfbr2)"
/note="vector: pAMP1; Site_1: NotI; Site_2: SalI"
```

ORIGIN

```
Query Match 17.8%; Score 586.4; DB 1; Length 589;
Best Local Similarity 99.7%; Pred. No. 2.5e-161;
Matches 587; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 517 TATTTCCTTGATGTCTCTGACCAAGACACCCATCAACAGATATATTTCTCCCATGAAGACG 576
|
Db 589 TATTTCCTTGATGTCTCTGACCAAGACACCCATCAACAGATATATTTCTCCCATGAAGACG 530
|
QY 577 GTACGCGCATCGGGCGTGGAGCATCTGCTCGCGGTTCAGGTCACCAAGCAAAATCGCGCTGTTAG 636
|
Db 529 GTACGCGCATCGGGCGTGGAGCATCTGCTCGCGGTTCAGGTCACCAAGCAAAATCGCGCTGTTAG 470
|
QY 637 CGGGCCCATTAAGTTCTGCTCGGGCGGTCTGCGTCTGGCTGGCATAAATATCTCA 696
|
Db 469 CGGGCCCATTAAGTTCTGCTCGGGCGGTCTGCGTCTGGCTGGCATAAATATCTCA 410
|
QY 697 CTGCAATCAAAATTCAGCCGATAGCGGAACGGAAGCGACTGGAGTGCCATGTCGCGTT 756
|
Db 409 CTGCAATCAAAATTCAGCCGATAGCGGAACGGAAGCGACTGGAGTGCCATGTCGCGTT 350
|
QY 757 TTCAACAAACCATGCAAAATGCTGAATGAGGCGCATGTTCCCACTGCGATGCTGTTGCCA 816
|
Db 349 TTCAACAAACCATGCAAAATGCTGAATGAGGCGCATGTTCCCACTGCGATGCTGTTGCCA 290
|
QY 817 ACGATCAGATCGGCGTGGGCGCAATGCGGCGCAATTTACCGAGTCCGGGCTGGCGTGGTG 876
|
Db 289 ACGATCAGATCGGCGTGGGCGCAATTTACCGAGTCCGGGCTGGCGTGGTG 230
|
QY 877 CGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCGGT 936
|
Db 229 CGGATATCTCGGTAGTGGGATACGACGATACCGAAGACAGCTCATGTTATATCCCGCGGT 170
|
QY 937 TAACCAACCATCAAAACAGGATTTTCGCTGCTGGGCGCAAAACAGCGCTGGACCGCTTGTCTGC 996
|
Db 169 TAACCAACCATCAAAACAGGATTTTCGCTGCTGGGCGCAAAACAGCGCTGGACCGCTTGTCTGC 110
|
QY 997 AACTCTCTCAGGGCGAGCGGTGAAGGGCAATCAGCTGTTGCCCGCTCTCACTGTGTGAAAA 1056
|
Db 109 AACTCTCTCAGGGCGAGCGGTGAAGGGCAATCAGCTGTTGCCCGCTCTCACTGTGTGAAAA 50
|
QY 1057 GAAAAACCAACCTCGCGCCCAATACGCAAAACCGGCTCTCCCGCGCGTT 1105
|
Db 49 GAAAAACCAACCTCGCGCCCAATACGCAAAACCGGCTCTCCCGCGCGTT 1
```

Search completed: August 3, 2005, 02:02:52
Job time : 6745.2 secs

THIS PAGE BLANK (USPTO)